

Email: cyapbs-r@mail.nih.gov
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:1393813
 High quality sequence stop: 439.
 Location/Qualifiers

FEATURES

source

1..587
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:3513237"
 /clone_lib="Soares_NMAX_maxillary_process"
 /tissue_type="maxillary process"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTACCAATCTGAAGTGGAGGCGCGCCGCTTTTTTTTTTTTTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 138 a 160 c 121 g 168 t

ORIGIN

alignment_scores:

Quality: 553.00 Length: 117
 Ratio: 4.982 Gaps: 1
 Percent Similarity: 94.872 Percent Identity: 89.744

alignment_block:

US-09-327-750d-33 x BG228077/rev ..

Align seg 1/1 to reverse of: BG228077 from: 1 to: 587

13 MetGluAsnAspHisGlnLysGluLysGluLysGluLysProGlnAs 29
 |||:|||||
 587 ATGAAGATGACCATCAGAAAAAGCAGGAGGAAGAATAACCCACAAGA 538
 29 pThrIleLysArgGluPro.ValValAlaProThrPheGluAlaGlyLys 45
 |||:|||||
 537 TACCATCAGAAGGAGGAGCGGTGTGGCCCTGACCTCCAGGCTGGCACA 488
 46 AsnCysAlaProArgGlyClyArgArgPheArgValArgGlnProIl 62
 |||:|||||
 487 AACTGTGCACCTAGAGAGGTCGACGCGGTTCGCGGTCGACGCCCAT 438
 62 eSerHisTyrArgTrpAspLeuMetHisArgValGlyGluProGlnGlyA 79
 |||:|||||
 437 CGCTCACTATAGATGGGACCTGATGCAGAGGTTGGGGAGCCCCAGGAA 388
 79 rgMetArgGluGluAsnValGlnArgPheGlyGluAspMetArgGlnLeu 95
 |||:|||||
 387 GGATGAGAGAGGAGACGCTACAGAGGTTTGGGGGTGATGTGAGACAGCTC 338
 96 MetGluLysLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSe 112
 |||:|||||
 337 ATGGAGAAGCTGAGGAAGGCGAGCTGACCCACAGCCCTCGCGGCGGTAG 288
 112 rThrAspProProHisHisAspHisHisAspGluPheCysLeuMetPro 128
 |||:|||||
 287 CACTGACCCGCTCATCATGACCAACCATGATGAGTTTTCCTCATGCC 239

seq_name: gb_est2:BF608209

seq_documentation_block:

LOCUS BF608209 650 bp mRNA EST 01-APR-2001
 DEFINITION MVL_001088 Mouse 9-day fetus cDNA library ICRFP522 Mus musculus
 cDNA clone ICRFP522C0878 5', mRNA sequence.
 ACCESSION BF608209
 VERSION BF608209.1 GI:13504638
 KEYWORDS EST.

house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

1 (bases 1 to 650)
 Yahyawi, M., Hennig, S., Neidhardt, L., Radelof, U., Hermann, B.G.,
 Lehrach, H. and O'Brien, J.

TITLE

Detection of a high number of novel genes in a 9-day mouse embryo
 cDNA library normalised by oligonucleotide fingerprinting

JOURNAL

COMMENT

Contact: Hennig S
 laboraty 123, dept. Lehrach
 Max-Planck-Institut fuer Molekulare Genetik
 Innestr. 63-73, D-14195 Berlin, Germany
 Tel: +49 30 8413 1612
 Fax: +49 30 8413 1380
 Email: hennig@molgen.mpg.de

EST's are made from clones being representatives of clone clusters.
 Clone clusters were calculated from oligonucleotide fingerprints.

PCR Primers

FORWARD: 5'-GAGCTATTCCAGAAGTAGTGA-3'

BACKWARD: 5'-TAATGAGCTCACTATAGGG-3'

Seq primer: 5'-ATTAGGTGACACTATAG-3'

High quality sequence stop: 650.

FEATURES

source

1..650
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="ICRFP522C0878"
 /clone_lib="Mouse 9-day fetus cDNA library ICRFP522"
 /tissue_type="whole embryo"
 /dev_stage="embryonic 9-day"
 /lab_host="E.coli, XL1 blue"
 /note="Vector: pSVSport1; Site_1: NotI; Site_2: SalI;
 Library preparation by oligo_dT priming of RNA. Clones can
 be ordered from the Resource Center in Berlin,
 http://www.rzpd.de."

BASE COUNT 164 a 167 c 180 g 128 t 11 others

ORIGIN

alignment_scores:

Quality: 516.00 Length: 128
 Ratio: 4.526 Gaps: 0
 Percent Similarity: 89.062 Percent Identity: 75.781

alignment_block:

US-09-327-750d-33 x BF608209 ..

Align seg 1/1 to: BF608209 from: 1 to: 650

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHi 17
 |||:|||||
 206 ATGGGACTCCAAAGATCAAGCGGTGAAAAATCTCACATGGACACGACCA 255
 17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleLysArg 34
 |||:|||||
 256 TCACAAAAGAGGACGACGAGGACAAAGCCACCAAGATACCATCTTAATGG 305
 34 luProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
 |||:|||||
 306 ACCGAGCTGTGGCCCTTACCTCTAGGCTGGCAAAACTGTGCACCTTGA 355
 51 GlyLysArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67
 |||:|||||
 356 GNAGGTGCGAGCGGTTCCGGGTTCCGACGCCCATCGCTTACTATANATC 405
 67 pAspLeuMetHisArgValGlyClyProGlnGlyArgMetArgGluGluA 84
 |||:|||||
 406 GGACCTGATCCACAGCGGTTGGGAGGCCCATGGAGGATCAGAGACGACA 455
 84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
 |||:|||||

```

456 ACGTACAGAGGCTCTGGGGTGATGTGAGACAGCTCATGGAGAANCCTCAGG 505
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProHis 117
|||||
506 GAAAGCAGCTTAGCCACAGCTTGGGGGGTGTAGCACTGACNCGCCTCA 555
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
556 TCATGACTACCATGATTAGTTTGGCTCATGCC 589
seq_name: gb_est1:BE291071
seq_documentation_block:
LOCUS BE291071 583 bp mRNA EST 13-JUL-2000
DEFINITION 601086311F1_NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3500522 5',
mRNA sequence.
ACCESSION BE291071
VERSION BE291071.1 GI:9172545
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 583)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
DNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM8559 row: e column: 03
High quality sequence start: 5
High quality sequence stop: 503.
FEATURES
Location/Qualifiers
1..583
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3500522"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/organism="Mus musculus"
/strain="PCMV-SPORT6; Site_1: Sali;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT 165 a 136 c 201 g 81 t
ORIGIN
alignment_scores:
Quality: 505.50 Length: 120
Ratio: 4.554 Gaps: 2
Percent Similarity: 92.500 Percent Identity: 83.333
alignment_block:
US-09-327-750D-33 x BE291071 ..
Align seg 1/1 to: BE291071 from: 1 to: 583
1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHis 17
|||||
208 ATGGAGTCCAAAGATCAAGCGGTGAATCTCAACATCGGAGAAATGACCA 257
17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleLysArgG 34

```

```

|||||
258 TCAGAAAAAGGAGGAGGAGAAAGCCACAAGATACCATCAGAAGG 307
34 luProValValalaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
|||||
308 AGCCAGCTGTGGCCTGACCTCGAGGCTGGCAAAACCTGTGCACCTAGA 357
51 GlyGlyArgArgPheArgValArgGlnProIleSerHisTyArgTr 67
|||||
358 GGAGGTGCGAGGCGGTTCGGGGTTCGGAGCCCATCGCTCATATAGATG 407
67 pAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluGluA 84
|||||
408 GGACCTGTATGCAGAGGTTGGGAGGCCCCAGTCAGAGTATGAGAGGACA 457
84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
|||||
458 ACCTACAGAGGTTAGGGGCTGATGTGAGACAGCTCATGGAAGCTGAGG 507
101 GluArgGlnLeuSerHis...SerLeuArgAlaVal.SerThrAspProp 116
|||||
508 GAAGGCTGTGAGCCACAAGACATGCGAGGCGGTTAGACAATGACCCGA 557
116 roHisHis 118
|||||
558 CTCATCAT 565
seq_name: gb_gss:AZ936393
seq_documentation_block:
LOCUS AZ936393 637 bp DNA GSS 26-APR-2001
DEFINITION 2M0193L05F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
Clone UUGC2M0193L05 F, DNA sequence.
ACCESSION AZ936393
VERSION AZ936393.1 GI:13794974
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 637)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: Q193 row: L column: 05
Seq primer: CGTGTAAACAGCAGCGCCAGT
Class: plasmid ends
High quality sequence stop: 637.
FEATURES
Location/Qualifiers
1..637
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0193L05"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, Tl-resistant, F-"
/note="vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource

```


(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWP42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid pL1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

BASE COUNT 131 a 137 c 132 g 176 t 1 others
ORIGIN

alignment_scores:
Quality: 493.50 Length: 117
Ratio: 4.569 Gaps: 1
Percent Similarity: 92.308 Percent Identity: 84.615

alignment_block:
US-09-327-750D-33 x AZ936393/rev ..

Align seg 1/1 to reverse of: AZ936393 from: 1 to: 637

1 MetGluSerLys...AspGlnGlyAlaLysAsnLeuAsnMetGluAsnAs 16
|||||
382 ATGGAGTCCAAAGTGAACAGCGCTGAAATCTCAACATGGAATGA 333

16 pHisGlnLysLysGluGluLysGluGluLysProGlnAspThrLeLysA 33

332 CCATCAGGAAAGAGGAGAAAGAGAAAGCCACAGGATGTAGCAAAA 283

33 rgGluProValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49

282 GGGATCCGATGTGGCCCTGCTTCGAGCTGGAGACTACTACGTGCCT 233

50 ArgGlyArgArgArgPheArgValArgGlnProIleSerHisTyrAr 66

232 AGAGGAGGTGCGAGCGGTTCGGGTTCCGACGCCCATCGTCACTACAG 183

66 gTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGlu 83

182 ATGGGACCTGATGCATAGGGTTGGGGAGCCGCCAGGAGGATGAGAGG 133

83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99

132 AGAAGCTACAGAGGTTGGGGATGATGTGAGACAGCTCATGGAGAAGCTN 83

100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116

82 GG.GAAGGACCTGAGCCACAGCCTCGGGGGGTAGCACTGACCCGCC 34

116 o 116

33 T 33

seq_name: gb_est2:BG669326

seq_documentation_block:

LOCUS BG669326 412 bp mRNA EST 30-APR-2001
DEFINITION DRNAEB11 Rat DRG Library Rattus norvegicus cDNA clone DRNAEB11

5', mRNA sequence.

ACCESSION BG669326 GI:13891248

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 412)

Xiao.H.S., Han,Z.G., Zhang,F.X., Huang,Q.H., Lu,Y.J., Bao,L., Fu,G., Guo,C., Yan,Q., Jin,S.X., Zhu,Z.D., Xu,X.R., Li,N.G., Chen,Z. and Zhang,X.

Distinct gene expression profiles of rat dorsal root ganglion induced by peripheral nerve axotomy

Unpublished (2001)

Contact: Zhang Xu

Laboratory of Sensory System

Institute of Neuroscience

320 Yue Yang Road, Shanghai 200031, P.R.China

Tel: 86-21-64748700-121

Fax: 86-21-64713446

Email: xu.zhang@ion.ac.cn

This clone is also available at Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong New Area, P.R.China. Please contact with Zhang Xu

(xu.zhang@ion.ac.cn) or Han Zequang (hanzg@chgc.sh.cn)

PCR Primers

FORWARD: T3

BACKWARD: T7

Seq primer: T3

POLYA-No.

Location/Qualifiers

1..412

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="DRNAEB11"

/clone_lib="Rat DRG Library"

/sex="male"

/tissue_type="dorsal root ganglion"

/dev_stage="adult"

BASE COUNT 124 a 83 c 130 g 75 t

ORIGIN

alignment_scores:

Quality: 489.50 Length: 126

Ratio: 4.450 Gaps: 2

Percent Similarity: 87.302 Percent Identity: 76.984

alignment_block:

US-09-327-750D-33 x BG669326 ..

Align seg 1/1 to: BG669326 from: 1 to: 412

1 MetGluSerLys...AspGlnGlyAlaLysAsnLeuAsnMetGluAsnAs 16

37 ATGGAGTCCAAAGTGAACAGCGGTGAAATCTCAACATGGAATGA 86

16 pHisGlnLysLysGluGluLysGluGluLysProGlnAspThrLeLysA 33

87 CCATCAGGAAAGAGGAGAAAGAGAAAGCCGACGATGCTTACAAA 136

33 rgGluProValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49

137 GGGAGCGGTTGTAGCCCTGCCTTTTGAAGCTGGAGAATACTACGTGCCT 186

50 ArgGlyArgArgArgPheArgValArgGlnProIleSerHisTyrAr 66

187 ATAGGAAGTCGAAGCGGTTCGCGTTTGGCAGCCCATTTGCCACTATAT 236

66 g.TrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGlu 82

237 GATGGACCTGATGCACAGGTTGGAGAGCCCAAGCAGGATGAGAGAA 286

83 GluAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLe 99

287 GAGATGTCAAAGGTTTGGGGAGGATATGAGACAGCTCATGTGAAAT 336

Percent Similarity: 89.231 Percent Identity: 70.000
 alignment_block:
 US-09-327-750D-33 x BI183527/rev ..

Align seg 1/1 to reverse of: BI183527 from: 1 to: 701

```

1 MetGluSerLysAsp.....GlnGlyAlaLysAsnLeuAsnMetGluAs 15
||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
653 ATGGCGTCCAAAGAGGAGCAGCAGTGAATAATCTCAGCATGGAAAA 604

15 nAspHisGlnLysLysGluGluLysGluGluLysProGlnAspThrIleL 32
| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
603 TACCACCAAGGAGAAATGAAAAAAGAACAAAAAGGGGCAAGATGCTAATA 554

32 ysArgGluProValValAlaProThrPheGluAlaGlyLysAsnCysAla 48
|| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
553 AAGGAGAGCCTTTGGCCCTCCCTGTG...GAAGCTGTGATATTCGCTGA 507

49 ProArgGlyGlyValArgArgPheArgValArgGlnProIleSerHisT 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
506 CCTAGAGGAATCGTAGGCGGTTCGTGTGAGCAGCCCTATCTCGCAGTA 457

65 rArgTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgG 82
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
456 TAGATGGGATATGACGCAGAGCTTGGACAGCCACAGGCAAGGATGAGAG 407

82 GluAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLys 98
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
406 AAGAGAATATGGAAGGATTGGGGAGGAGGTGAGGCTGCTGATGGAAGAAG 357

99 LeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPr 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
356 CTGAGGGAAACACGATTGAGTCATAGTCTTCGGGAGTTAGCAGCTGACCC 307

115 OProHisAspHisHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
306 CCCTCACCATGACCATCAGGATGAGTTTGCCTTATGCCT 267

```

seq_name: gb_est1:AW536634

```

seq_documentation_block:
LOCUS AW536634 549 bp mRNA EST 31-AUG-2000
DEFINITION G0106H08-3 NIA Mouse E7.5 Embryonic Portion cDNA Library Mus
musculus cDNA clone G0106H08 3', mRNA sequence.
ACCESSION AW536634
VERSION AW536634.1 GI:7179051
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 549)
AUTHORS Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac
,M.J., Pantano,S., Sano,Y., Piao,Y., Negaraja,R., Dol,H., Wood,W.H.
,III, Becker,K.G. and Ko,M.S.H.
TITLE Genome-wide expression profiling of mid-gestation placenta and
embryo using a 15,000 mouse developmental cDNA microarray
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
MEDLINE 20381348
COMMENT Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: G0106 row: H column: 08
Seq primer: -21M13 Forward
High quality sequence stop: 549
POLYA=Yes.
FEATURES
source Location/Qualifiers
1..549
/organism="Mus musculus"
/strain="C57BL/6J"

```

```

99 uArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProp 116
| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
337 GTGGAGAGGACACTGTGCCATAATCTGCGGGCGGTGAGCACTGTCCCCC 386

116 roHisHisAspHisHisAspGluPhe 124
| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
387 CTTTACCATGACCACCATGAGTGT 412

seq_name: gb_est2:BI183527

seq_documentation_block:
LOCUS BI183527 701 bp mRNA EST 10-JUL-2001
DEFINITION UNL-P-FN-by-f-07-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
UNL-P-FN-by-f-07-0-UNL 3', mRNA sequence.
ACCESSION BI183527
VERSION BI183527.1 GI:14657936
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 701)
Caetano,A.R., Johnson,R.K. and Pomp,D.
Generation and sequence characterization of a normalized cDNA
library from swine ovarian follicles
Unpublished (2001)
Contact: Pomp, D
Department of Animal Science
University of Nebraska, Lincoln
Lincoln, NE 68583-0908, USA
Tel: 402 472 6416
Fax: 402 472 6362
Email: dpomp@unl.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
normalized porcine ovarian follicles library
Seq primer: M13 -29
POLYA=Yes.

```

```

FEATURES
source Location/Qualifiers
1..701
/organism="Sus scrofa"
/strain="University of Nebraska, Lincoln Swine Selection
Lines"
/db_xref="taxon:9823"
/clone="UNL-P-FN-by-f-07-0-UNL"
/clone_lib="UNL-P-FN"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/notes="vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UNL-P-FN
library is a normalized library representing porcine
ovarian follicles, ranging between 2.0 to 10.0 mm in
diameter, collected during 7 days of the follicular phase
of the pig estrous cycle. This library was derived from
the library UNL-P-F2. The tag is a string of 5-6
nucleotides present between the Not I site and the
oligo-dT track. The library was constructed as described
by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
, 1996
TAG_LIB=UNL-P-FN
TAG_TISSUE=porcine ovarian follicles
TAG_SEQ=CACACT"
BASE COUNT 173 a 190 c 119 g 218 t 1 others
ORIGIN
alignment_scores:
Quality: 480.50 Length: 130
Ratio: 4.142 Caps: 2

```

THIS PAGE BLANK (USPTO)

```
/db_xref="taxon:10090"
/clone="G0106H08"
/clone_lib="NIA Mouse E7.5 Embryonic Portion cDNA Library"
/sex="unknown"
/dev_stage="7.5dpc Embryo"
/lab_host="DH10B"
/Note="Vector: pSPORT1 (Gibco/BRL Life Technology);
Site 1: SalI; Site 2: NotI; Total RNAs were extracted from
6 Embryo. The double-stranded cDNA was synthesized by
Gibco's kit with an oligo(dT) primer (NotI primer-adaptor
from GibcoBRL)
[5'-pGACTAGTTCTAGATCGAGCGCCGCTTTTTTTTTTTT-3']
from 0.5ug of mRNA. The double-stranded cDNAs were
treated with T4 DNA polymerase and purified by
ethanol-precipitation. The cDNAs were ligated to
Lone-linker L1-Sal3 (include SalI sequence). The cDNAs
were purified by phenol/chloroform and separated from
free linkers by Centricon 100. Then, cDNAs were amplified
by long-range high fidelity PCR using Takara's Ex Taq
polymerase. Then, the cDNAs were purified by
phenol/chloroform and by Centricon 100. The cDNAs were
digested with SalI and NotI enzymes. Then, the cDNAs were
size selected by Gibco's Size Fractionation Column. The
cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by chemical method. The library was
constructed by Xiaohong Wang and Minoru S. H. Ko."
```

BASE COUNT 151 a 151 c 117 g 130 t
ORIGIN

alignment_scores:
Quality: 472.00 Length: 97
Ratio: 5.187 Gaps: 0
Percent Similarity: 93.814 Percent Identity: 88.660

alignment_block:

US-09-327-750D-33 x AW536634/rev ..

Align seg 1/1 to reverse of: AW536634 from: 1 to: 549

```
32 LysArgGluProValValAlaProThrPheGluAlaGlyLysAsnCysAl 48
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
541 AAAAGGGATCCGATTCTGCCCTTCCCTTCGAAGCTGGAGACTACTACGT 492
48 aProArgGlyGlyArgArgPheArgValArgGlnProIleSerHist 65
:|||||:|||||:|||||:|||||:|||||:|||||:|||||
491 GCCTAGAGGAGGTCGACGCGGTTCGGGTTCGGACGCCCATCGTGCAC 442
65 YrArgTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArg 81
|||||:|||||:|||||:|||||:|||||:|||||:|||||
441 ACAGATGGACCTGATGCTAGGTTGGGAGCCCCCGGAGGAGGTGAGA 392
82 GluGluAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLy 98
|||||:|||||:|||||:|||||:|||||:|||||:|||||
391 GAGGACAACGTACAGAGTTTGGGGATGATGTCAGACACCTCATCGAGAA 342
98 sLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAsp 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||
341 GCTGAGGGAAGGACACTGAGCCACACCTGCGGGCGGTAGCACCTGACC 292
115 rGProHisHisAspHisHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||
291 CGCCTCATCATGACCAACCATGATGAGTTTGGCTTTATGCC 251
```

OM of: US-09-327-750D-33 to: Issued_Patents_NA:* out_format : pfs

Date: Mar 11, 2002 3:51 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frameat_p2n.model -DEV=xlh
-O/cgn2_1/USPTO.spool/US09327750/runat_11032002_101154_20340/app_query.fasta_1.1472
-DB=Issued_Patents_NA -OPMT=fastat -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOPEXT=0.000
-GAPOP=4.500 -GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
-FCAP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blom62
-DELOP=6.000 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-TRANS=human40.cdi -LIST=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-THR_MAX=100 -THR_MIN=0 -MINLEN=0 -MAXLEN=2000000000
-NORM=ext -HEAPSZE=500 -MINLEN=0 -ICPU=3 -LONGLOG
-USR=US09327750.@CGL1_1_99 -NCP=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-33

Query length: 128

Database: Issued_Patents_NA:*

Database sequences: 351203

Database length: 113238999

Search time (sec): 146.090000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-381-881-1 +	93.50	166.92	0.1200	1315	1
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-281-221-1 +	93.50	166.92	0.1200	1315	1
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-07-951-715A-6 +	93.00	155.63	0.5107	3624	1
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-459-448A-6 +	93.00	155.63	0.5107	3624	1
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-459-595A-6 +	93.00	155.63	0.5107	3624	1
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-459-504B-6 +	93.00	155.63	0.5107	3624	1
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-459-444-6 +	93.00	155.63	0.5107	3624	1
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-053-549-1 +	93.00	155.63	0.5107	3624	1
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-053-549-1 +	91.00	142.54	2.74	8854	1
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-506-553C-6 +	82.50	152.65	0.7483	607	1
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-998-416-994 -	82.00	149.79	1.08	729	1
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-426-599B-3 +	81.00	149.35	1.14	624	1
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-147-777-3 +	80.50	110.90	158.26	24979	1
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-452-872-3 +	80.50	110.90	158.26	24979	1
/cgn2_6/ptodata/2/ina/PTCTUS.COMB.seq:PCT-US93-03985-3 +	80.50	110.90	158.26	24979	1
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-434-288-10 +	80.00	137.79	5.03	1600	1
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-030-995-1 +	79.00	137.39	5.30	1365	1
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-117-927-4 -	78.50	131.71	10.97	2163	1
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-931-999-4 +	77.50	118.14	62.53	6755	1
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-125-468-1 -	77.50	103.01	435.42	30001	1
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-474-933-1 -	77.50	103.01	435.42	30001	1
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-08-998-416-141 -	77.00	140.56	3.53	672	1
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-047-041A-24 -	77.00	133.96	8.23	1288	1
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-285-440-11 +	76.50	130.76	12.40	1599	1
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-630-349-11 +	76.50	130.76	12.40	1599	1
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-285-440-13 +	76.50	127.81	18.09	2137	1
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-630-349-13 +	76.50	127.81	18.09	2137	1
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-306-691B-23 +	76.50	127.06	19.92	2301	1
/cgn2_6/ptodata/2/ina/PTCTUS.COMB.seq:PCT-US93-06251-78 +	76.50	127.06	19.92	2301	1
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-056-200-93 +	76.50	112.62	127.02	9551	1
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-800-644-93 +	76.50	112.62	127.02	9551	1
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-09-036-987A-1 -	76.00	90.02	2.3e+03	80161	1
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-370-700-1 -	76.00	90.02	2.3e+03	80161	1
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-717-294-42 +	75.50	118.36	60.85	4451	1
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-103-840A-2 +	75.50	48.36	3.5e+05	4403765	1
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-381-881-3 +	75.00	137.29	5.37	624	1
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-281-221-3 +	75.00	137.29	5.37	624	1
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-809-740A-1 -	75.00	120.54	46.01	3252	1
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-809-740A-4 -	75.00	120.54	46.01	3252	1
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-822-586-45 +	75.00	109.04	201.00	10095	1

/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-452-075-1 - 74.50 120.90 43.90 2841
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-231-061-1 - 74.50 120.90 43.90 2841
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-403-852D-1 + 74.50 114.40 101.08 5392
/cgn2_6/ptodata/2/ina/6A.COMB.seq:US-08-510-646B-1 + 74.50 114.40 101.08 5392
/cgn2_6/ptodata/2/ina/6B.COMB.seq:US-09-231-818-1 + 74.50 114.40 101.08 5392

seq_name: /cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-381-881-1

seq_documentation_block:
; Sequence 1, Application US/08381881
; Patent No. 5939064
; GENERAL INFORMATION:
; APPLICANT: SAVALKOU, PAUL H., M.
; APPLICANT: CAASTRA, WILLEM
; TITLE OF INVENTION: BORDETTELLA BRONCHISEPTICA VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVENS, DAVIS, MILLER & MOSHER
; STREET: 515 NORTH WASHINGTON STREET
; CITY: ALEXANDRIA
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381.881
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: POULOS III, JAMES A.
; REGISTRATION NUMBER: 31,714
; REFERENCE/DOCKET NUMBER: TPP 29685
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703 549-7200
; TELEFAX: 703 528-5313
; TELEX: 44-0704 "STEVENS"
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1315 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bordetella bronchiseptica
; STRAIN: 401
; IMMEDIATE SOURCE:
; CLONE: E coli PC2495(pIVB3-420)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..539
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 540..1142
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1143..1315
; US-08-381-881-1

alignment_scores:
Quality: 93.50 Length: 96
Ratio: 1.948 Gaps: 4
Percent Similarity: 50.000 Percent Identity: 33.333
alignment_block:
US-09-327-750D-33 x US-08-381-881-1

```
Align seg 1/1 to: US-08-381-881-1 from: 1 to: 1315
43 AlaGlyLysAsnCysAlaProArgGlyGlyArgArgPheArgValAr 59
||||:||||| ||| ||| ||| ||| ||| ||| |||
542 GCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 591
59 gGlnProIle.....SerHisTyrArgTrrp..... 67
||||| ||| ||| ||| ||| ||| ||| ||| |||
592 CCGCCCATGCCGAGAGCGGACCATGTCTACCGCGCAGCATCACCGAC 641
68 ..AspLeuMethHisArgValGlyGluPro.....GlnGlyArgMet 80
||||| ||| ||| ||| ||| ||| ||| ||| |||
642 CAGACCTGCAGCATCGAGGACCGCGCGGTTACATCAAGGTCGTGCA 691
81 ArgGluGluAsnValGlnArgPheGlyGlyAspMetArgGlnLeuMetGl 97
||||| ||| ||| ||| ||| ||| ||| ||| |||
692 CCT.....GCCACGATCTCCAGAGCGGCGCTGA 720
97 uLysLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThra 114
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
721 AGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 770
114 spProProHisHisAspHisAspPheCysLeu 126
|||| ||| ||| ||| ||| ||| ||| ||| |||
771 GACTGCCGCGACCGCGTCAACACTCTCAAGCTGTACTT 808

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-281-221-1

seq_documentation_block:
; Sequence 1, Application US/09281221
; Patent No. 6284256
; GENERAL INFORMATION:
; APPLICANT: SAELEKOU, PAUL H.M.
; GASTRA, WILLEM
; TITLE OF INVENTION: BORDETTELLA BRONCHISEPTICA VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVENS, DAVIS, MILLER & MOSHER
; STREET: 515 NORTH WASHINGTON STREET.
; CITY: ALEXANDRIA
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/281,221
; FILING DATE: 30-Mar-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/381,881
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: POULOS III, JAMES A.
; REGISTRATION NUMBER: 31,714
; REFERENCE/DOCKET NUMBER: TPP 29685
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703 549-7200
; TELEFAX: 703 528-5313
; TELEX: 44-0704 "STEVENS"
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1315 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

```
; ORIGINAL SOURCE:
; ORGANISM: Bordetella bronchiseptica
; STRAIN: 401
; IMMEDIATE SOURCE:
; CLONE: E coli PC2495(pIVB3-420)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..539
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 540..1142
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1143..1315
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-281-221-1

alignment_scores:
Quality: 93.50 Length: 96
Ratio: 1.948 Gaps: 4
Percent Similarity: 50.000 Percent Identity: 33.333

alignment_block:
US-09-327-750D-33 x US-09-281-221-1
Align seg 1/1 to: US-09-281-221-1 from: 1 to: 1315
43 AlaGlyLysAsnCysAlaProArgGlyGlyArgArgPheArgValAr 59
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
542 GCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 591
59 gGlnProIle.....SerHisTyrArgTrrp..... 67
||||| ||| ||| ||| ||| ||| ||| ||| |||
592 CCGCCCATGCCGAGAGCGGACCATGTCTACCGCGCAGCATCACCGAC 641
68 ..AspLeuMethHisArgValGlyGluPro.....GlnGlyArgMet 80
||||| ||| ||| ||| ||| ||| ||| ||| |||
642 CAGACCTGCAGCATCGAGGACCGCGCGGTTACATCAAGGTCGTGCA 691
81 ArgGluGluAsnValGlnArgPheGlyGlyAspMetArgGlnLeuMetGl 97
||||| ||| ||| ||| ||| ||| ||| ||| |||
692 CCT.....GCCACGATCTCCAGAGCGGCGCTGA 720
97 uLysLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThra 114
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
721 AGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 770
114 spProProHisHisAspHisAspPheCysLeu 126
|||| ||| ||| ||| ||| ||| ||| ||| |||
771 GACTGCCGCGACCGCGTCAACACTCTCAAGCTGTACTT 808

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-951-715A-6

seq_documentation_block:
; Sequence 6, Application US/07951715A
; Patent No. 5625136
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
```

```

; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/951,715A
; FILING DATE: 25-SEP-1992
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP
; TELEPHONE: (919)541-8615
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; OTHER INFORMATION: optimized cryiB"
; OTHER INFORMATION: /note= "Disclosed in Figure 6."
;
; US-07-951-715A-6

```

```

alignment_scores:
  Quality: 93.00      Length: 116
  Ratio: 1.550       Gaps: 4
  Percent Similarity: 51.724   Percent Identity: 26.724

```

```
alignment_block:
US-09-327-750D-33 x US-07-951-715A-6
..
```

```
Align seg 1/1 to: US-07-951-715A-6 from: 1 to: 3624
```

```

14  GluAsnAspHisGlnLysLysGluLysGluLysProGlnAspTh 30
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1026 GAACACGAGCCACCCAGCGCGCCACACACGAGCATCAACCCCGTGACCC 1075

30  rileLysArgGluProValAlaProThrPheGluAlaGlyLysAsnC 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1076 TGCCTTTCGCGAGCGCGAGCGTGTCACCGCAGAGCATCGCGCGGTG 1125

47  ysAla.....ProArgGlyGlyArgArg 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1126 CTGCTGTGGCGCATCTACCTGGAGCCCATCAACGCGTGCACCGGTGCG 1175

56  PheArgValArgGlnProIleSerHisTyrArgTrpAspLeuMetHisAr 72
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1176 CTTCAACTTCAACCAACCCCGGACACATCAGCGA.....CCGCGGCACCG 1219

```

```

72  gValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnArgPheG 89
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1220 CCAACTACAGCCA.....GCCCTACGAGAGCCCGCGCTGCAGCTGAAG 1263

89  lyGluAspMetArgGlnLeuMetGluLysLeuArgGlnLeuSer 105
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1264 GACAGCGAGACCGAGCTGCCCGCCGAGACGACGAGCGGCCCAA..... 1307

106  HisSerLeuArgAlaValSerThrAspProHisHisAspHis 121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1308 .....CTACGAGAGCTACAGCCACCGCTGAGCCACATCGGCATCAT 1349

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-459-448A-6

seq_documentation_block:
; Sequence 6, Application US/08459448A
; Patent No. 5859336
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NO. 5859336artis Corporation
; STREET: Patent & Trademark Dept., 520 White Plains
; STREET: Rd., POB 2005
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,448A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40403
; REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8582
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

```
;
;
; MOLECULE TYPE: linear
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
;
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; OTHER INFORMATION: optimized cryIB"
; OTHER INFORMATION: /note= "Disclosed in Figure 6."
;
US-08-459-448A-6

alignment_scores:
  Quality: 93.00 Length: 116
  Ratio: 1.550 Gaps: 4
  Percent Similarity: 51.724 Percent Identity: 26.724

alignment_block:
US-09-327-750D-33 x US-08-459-448A-6 ..
  Align seg 1/1 to: US-08-459-448A-6 from: 1 to: 3624

14 GluAsnAspHisGlnLysGluGluLysGluLysProGlnAspTh 30
|||||
1026 GAACACGACGACCCAGCGCGGCCAACACACGACGATCAACCCCGTGACCC 1075

30 rLeLysArgGluProValAlaProThrPheGluAlaGlyLysAsnC 47
|||||
1076 TGCCTTCGCGACGCGCGGACGTGTACCGCACCGAGAGCTACCGCGCGTG 1125

47 ysAla.....ProArgGlyGlyArgArg 55
|||||
1126 CTGCTGTGGGCGATCTACCTGGAGCCCATCCAGCGCTGCCACCGTGGC 1175

56 PheArgValArgGlnProIleSerHisTyArgTrpAspLeuMethHisAr 72
|||||
1176 CTTCAACTTCACCAACCCCGACACATCAGCGA.....CCGCGGCACCG 1219

72 gValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnArgPheG 89
|||||
1220 CCAACTACAGCA.....GCCCTACGAGAGCGCGCGCTGCAGCTGAAG 1263

89 lyGluAspMetArgGlnLeuMetGluLysLeuArgGluArgGlnLeuSer 105
|||||
1264 GACAGCGAGACGAGCTGCCCGCCGAGACACCGAGCGGCCCAA..... 1307

106 HisSerLeuArgAlaValSerThrAspProProHisHisAspHisHis 121
|||||
1308 .....CTAGGAGAGCTACAGCCACCGCTGAGCCACATCGGCATCAT 1349

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-459-595A-6

seq_documentation_block:
; Sequence 6, Application US/08459595A
; Patent No. 6018104
;
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Meriin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
```



```
56 PheArgValArgGlnProHisSerHisTyrArgTrpAspLeuMethHisAr 72
1176 CTTCACTTCCACCAACCCCGAGAACATCAGCGA.....CCGCGGCACCG 1219
72 gValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnArgPheG 89
1220 CCAACTACAGCCA.....GCCCTACAGAGAGCCCGCGCTGCAGCTGAAG 1263
89 lyGluAspMetArgGlnLeuMetGluLysLeuArgGluArgGlnLeuSer 105
1264 GACAGCCAGACCGAGTGGCCCGGAGACACCGAGCGGCCCA..... 1307
106 HisSerLeuArgAlaValSerThrAspProHisHisAspHisHis 121
1308 .....CTACGAGAGCTACAGCACCCTGAGCCACATCGGCATCAT 1349
```

seq_name: /cqn2_6/ptodata/2/ina/6A_COMB.seq:US-08-459-504B-6

```
seq_documentation_block:
; Sequence 6, Application US/08459504B
; Patent No. 6075185
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 6075185artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,504B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/459,595
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; OTHER INFORMATION: optimized cryIB"
; OTHER INFORMATION: /note= "Disclosed in Figure 6."
US-08-459-504B-6

alignment_scores:
Quality: 93.00 Length: 116
Ratio: 1.550 Gaps: 4
Percent Similarity: 51.724 Percent Identity: 26.724

alignment_block:
US-09-327-750D-33 x US-08-459-504B-6
Align seg 1/1 to: US-08-459-504B-6 from: 1 to: 3624

14 GluAsnAspHisGlnLysLysGluGluLysGluLysProGlnAspTh 30
||||:||||| :||||:||||| :||||:|||||
1026 GAACACGACGACCGCGCCGACCAACACGACGACGACGACGACG 1075
30 rleLysArgGluProValValAlaProThrPheGluAlaGlyLysAsnC 47
||||:||||| :||||:||||| :||||:|||||
1076 TGGCTTGGCCGCGGAGCTGTACCGACGAGCTAGCGCGCGGTG 1125
47 ysAla.....ProArgGlyGlyArgArg 55
||||:||||| :||||:|||||
1126 CTGCTGTGGGGCATCTACCTGGAGCCCATCCACGCGTGCACCGTGC 1175
56 PheArgValArgGlnProHisSerHisTyrArgTrpAspLeuMethHisAr 72
||||:||||| :||||:||||| :||||:|||||
1176 CTTCACTTCCACCAACCCCGAGAACATCAGCGA.....CCGCGGCACCG 1219
72 gValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnArgPheG 89
||||:||||| :||||:||||| :||||:|||||
1220 CCAACTACAGCCA.....GCCCTACAGAGAGCCCGCGCTGCAGCTGAAG 1263
89 lyGluAspMetArgGlnLeuMetGluLysLeuArgGluArgGlnLeuSer 105
||||:||||| :||||:||||| :||||:|||||
1264 GACAGCCAGACCGAGTGGCCCGGAGACACCGAGCGGCCCA..... 1307
106 HisSerLeuArgAlaValSerThrAspProHisHisAspHisHis 121
||||:||||| :||||:||||| :||||:|||||
1308 .....CTACGAGAGCTACAGCACCCTGAGCCACATCGGCATCAT 1349

seq_name: /cqn2_6/ptodata/2/ina/6A_COMB.seq:US-08-459-444-6
```

```
seq_documentation_block:
; Sequence 6, Application US/08459444A
; Patent No. 6121014
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED
; NUCLEIC ACID CODING SEQUENCE
```

```

;
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,444A
; FILING DATE: 02-Jun-1995
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-18805/PI/CGC1577/CIP/DIV6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; optimized cryiB"
; /note= "Disclosed in Figure 6."
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
;
; US-08-459-444-6
;
;
; alignment_scores:
; Quality: 93.00 Length: 116
; Ratio: 1.550 Gaps: 4
; Percent Similarity: 51.724 Percent Identity: 26.724
;
; alignment_block:
; US-09-327-750D-33 x US-08-459-444-6 ..
;
; Align seg 1/1 to: US-08-459-444-6 from: 1 to: 3624
;
; 14 GluAsnAspHisGlnLysLysGluGluLysProGlnAspTh 30
; |||:|||||:|||||:|||||:|||||:|||||:|||||:
; 1026 GAACACGAGCAGCCAGCGCGCCACCAACACGAGCTACCGCGGTGCC 1075
;
; 30 rLeLysArgGluProValAlaProThrPheGluAlaGlyLysAsnC 47
; |||:|||||:|||||:|||||:|||||:|||||:|||||:
; 1076 TGCCTTCGCGAGCGCGGACGTGTACCGCACCGAGAGCTACGCGCGGTG 1125
;
; 47 ysAla.....ProArgGlyGlyArgArgArg 55
; |||:|||||:|||||:|||||:|||||:|||||:|||||:
; 1126 CTGCTGTGGGCATCTACCTGGAGCCCATCCACGCGGTGCCACCGTGGC 1175
;
; 56 PheArgValArgGlnProIleSerHisTyrArgTrpAspLeuMethIsAr 72
; |||:|||||:|||||:|||||:|||||:|||||:|||||:
; 1176 CTTCAACTTCAACCAACCCCAAGAACATCAGCGA.....CCGCGGACCG 1219

```

```

72 gValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnArgPheG 89
; |||:|||||:|||||:|||||:|||||:|||||:|||||:
; 1220 CCAACTACACCA.....GCCCTACGAGAGAGCCCGCCCTGCAGCTGAAG 1263
;
; 89 lyGluAspMetArgGlnLeuMetGluLysLeuArgGluArgGlnLeuSer 105
; |||:|||||:|||||:|||||:|||||:|||||:|||||:
; 1264 GACAGCGAGACCGAGCTGCCCGCCGAGACCGAGCGGCCCA..... 1307
;
; 106 HisSerLeuArgAlaValSerThrAspProPheHisHisAspHisHis 121
; |||:|||||:|||||:|||||:|||||:|||||:|||||:
; 1308 .....CTAGGAGAGCTACAGCCACCGCCTGAGCCACATCGGCATCAT 1349
;
; seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:us-09-053-549-7
;
; seq_documentation_block:
; Sequence 7, Application US/09053549
; Patent No. 6121521
; GENERAL INFORMATION:
; APPLICANT: Desai, Nallni
; TITLE OF INVENTION: No. 6121521el Insecticidal Protein and Gene
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6121521artis Corporation
; STREET: 3054 Cornwallis Rd.
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,549
; FILING DATE: 01-APR-1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; optimized cryiB"
; OTHER INFORMATION: /note= "Disclosed in Figure 6."
;
; US-09-053-549-7
;
;
; alignment_scores:
; Quality: 93.00 Length: 116
; Ratio: 1.550 Gaps: 4
; Percent Similarity: 51.724 Percent Identity: 26.724
;
; alignment_block:
; US-09-327-750D-33 x US-09-053-549-7 ..
;
; Align seg 1/1 to: US-09-053-549-7 from: 1 to: 3624

```

```
14 GluAsnAspHisGlnLysLysGluGluLysGluLysProGlnAspTh 30
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1026 GAACACGAGCAGCCAGCGGCCACCAACAGCAGCATCAACCCCGTGACCC 1075
30 rleLysArgGluProValValAlaProThrPheGluAlaGlyLysAsnC 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1076 TGGCGTTTCGCGCAGCGCGAGCTGTACGCGACCGAGAGCTACGCGCGGTG 1125
47 ySala.....ProArgGlyGlyArgArgArg 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1126 CTGCTGTGGGGCATCTACCTGGAGCCCATCCAGCGGTGCCACCGTGGC 1175
56 PheArgValArgGlnProIleSerHisTyrArgTirpAspLeuMethHisAr 72
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1176 CTTCAACTTCACCAACCCCGCAGAACATCAGCGA.....CCGCGGCACCG 1219
72 gValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnArgPheG 89
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1220 CCAACTACAGCCA.....GCCCTACGAGAGAGCCCGCGCTGCAGCTGAAG 1263
89 lyGluAspMetArgGlnLeuMetGluLysLysLeuArgGluArgGlnLeuSer 105
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1264 GACAGCAGCAGCGAGCTGCCCGCGAGACACCGAGCGCCCA..... 1307
106 HisSerLeuArgAlaValSerThrAspProHisHisAspHisHis 121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1308 .....CTACGAGAGCTACAGCCAGCCGCTGAGCCACATCGGCATCAT 1349
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-053-549-1
```

```
seq_documentation_block:
; Sequence 1, Application US/09053549
; Patent No. 6121521
; GENERAL INFORMATION:
; APPLICANT: Desai, Nalini
; TITLE OF INVENTION: No. 6121521el Insecticidal Protein and Gene
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6121521artis Corporation
; STREET: 3054 Cornwallis Rd.
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,549
; FILING DATE: 01-APR-1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 11..3694
; OTHER INFORMATION: /product= "hyFLIB protein"
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: 8854
; OTHER INFORMATION: /note= "Sequence of pCIB5520
; Patent No. 6121521
; OTHER INFORMATION: containing coding sequence for hyFLIB protein"
US-09-053-549-1

alignment_scores:
Quality: 91.00 Length: 116
Ratio: 1.542 Gaps: 4
Percent Similarity: 50.862 Percent Identity: 27.586

alignment_block:
US-09-327-750D-33 x US-09-053-549-1 ..

Align seg 1/1 to: US-09-053-549-1 from: 1 to: 8854

14 GluAsnAspHisGlnLysLysGluGluLysGluLysProGlnAspTh 30
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1099 GAACACGAGCAGCCAGCGGCCACCAACAGCAGCATCAACCCCGTGACCC 1148
30 rleLysArgGluProValValAlaProThrPheGluAlaGlyLysAsnC 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1149 TGGCGTTTCGCGCTCCGAGAGCTGTACGCGACCGAGAGCTACGCGCGGTG 1198
47 ySala.....ProArgGlyGlyArgArgArg 55
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1199 CTGCTGTGGGGCATCTACCTGGAGCCCATCCAGCGGTGCCACCGTGGC 1248
56 PheArgValArgGlnProIleSerHisTyrArgTirpAspLeuMethHisAr 72
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1249 CTTCAACTTCACCAACCCCGCAGAACATCAGCGA.....CCGCGGCACCG 1292
72 gValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnArgPheG 89
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1293 CCAACTACAGCCA.....GCCCTACGAGAGCGCGGTTCAGCTGAAG 1336
89 lyGluAspMetArgGlnLeuMetGluLysLysLeuArgGluArgGlnLeuSer 105
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1337 GACAGCAGCAGCGAGCTGCCCGCGAGACACCGAGCGCCCA..... 1380
106 HisSerLeuArgAlaValSerThrAspProHisHisAspHisHis 121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1381 .....CTACGAGAGCTACAGCCAGCCGCTGAGCCACATCGGCATCAT 1422
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-506-553C-6

seq_documentation_block:
; Sequence 6, Application US/08506553C
; Patent No. 6120989
; GENERAL INFORMATION:
; APPLICANT: Vornhagen, Rolf; Hinderer, Walter; Sonnborn, Han-H.;
; APPLICANT: Plachter, Bodo; and Jahn, Gerhard
; TITLE OF INVENTION: ISOLATED HUMAN CYTOMEGALOVIRUS
; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felle & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/506,553C
; FILING DATE: 07/25/95
; CLASSIFICATION: 435
```

seq_name: /cqn2_6/ptodata/2/ina/6B_COMB.seq:US-08-998-416-994


```

alignment_scores:
  Quality: 80.50      Length: 72
  Ratio: 1.872       Gaps: 1
  Percent Similarity: 59.722  Percent Identity: 31.944

alignment_block:
US-09-327-750D-33 x US-08-147-777-3
..
Align seg 1/1 to: US-08-147-777-3 from: 1 to: 24979

50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
||||| |||||:|||||:|||||: |||
5088 CGCGGAGGCGCGGCAGCTACGGCTCAGGAGCTACGGCGGCAGCG 5137

66 gTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83
||||| ||| :||| ||||| |||
5138 GTGGGGGCTATGGCGCGGGAAGAGCTACCGCGGAGCGGCAGCAGGC 5187

83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
:||||:|||||: |||: |||||: |||||: |||||: |||||
5188 GGGAGTGGAGCGGGTTATGCGAGCGGCTGCGCGCGGTGGCGGAGCTA 5237

100 ...ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPr 115
||||| ||||| ||||| ||||| ||||| ||||| |||||
5238 CGGAGGAGCGGCAGAGAGCGCGCGGATCCTCGCGCGTCAGATCATCC 5287

115 oProHisHisAspHis 120
| |||||: |||||
5288 AGACCTCCACCAACAC 5303

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-452-872-3

seq_documentation_block:
; Sequence 3, Application US/08452872
; Patent No. 6057298
; GENERAL INFORMATION:
; APPLICANT: Roop, Dennis R.
; APPLICANT: Rothnagel, Joseph A.
; APPLICANT: Greenhalgh, David A.
; APPLICANT: Yuspa, Stuart H.
; TITLE OF INVENTION: KERATIN K1 EXPRESSION VECTORS
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: LYON & LYON
; STREET: 611 West Sixth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: IBM MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/452.872
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/147,777
; FILING DATE:
; APPLICATION NUMBER: 07/876,289
; FILING DATE: April 30, 1992
; APPLICATION NUMBER: Unassigned (204/144)
; FILING DATE: October 29, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 204/153
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24979 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-452-872-3

alignment_scores:
  Quality: 80.50      Length: 72
  Ratio: 1.872       Gaps: 1
  Percent Similarity: 59.722  Percent Identity: 31.944

alignment_block:
US-09-327-750D-33 x US-08-452-872-3
..
Align seg 1/1 to: US-08-452-872-3 from: 1 to: 24979

50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
||||| |||||:|||||:|||||: |||
5088 CGCGGAGGCGCGGCAGCTACGGCTCAGGAGCTACGGCGGCAGCG 5137

66 gTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83
||||| ||| :||| ||||| |||
5138 GTGGGGGCTATGGCGCGGGAAGAGCTACCGCGGAGCGGCAGCAGGC 5187

83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
:||||:|||||: |||: |||||: |||||: |||||: |||||
5188 GGGAGTGGAGCGGGTTATGCGAGCGGCTGCGCGCGGTGGCGGAGCTA 5237

100 ...ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPr 115
||||| ||||| ||||| ||||| ||||| ||||| |||||
5238 CGGAGGAGCGGCAGAGAGCGCGCGGATCCTCGCGCGTCAGATCATCC 5287

115 oProHisHisAspHis 120
| |||||: |||||
5288 AGACCTCCACCAACAC 5303

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US93-03985-3

seq_documentation_block:
; Sequence 3, Application PC/TUS9303985
; GENERAL INFORMATION:
; APPLICANT: Roop, Dennis R.
; APPLICANT: Rothnagel, Joseph A.
; APPLICANT: Greenhalgh, David A.
; APPLICANT: Yuspa, Stuart H.
; TITLE OF INVENTION: DEVELOPMENT OF A VECTOR TO TARGET GENE
; TITLE OF INVENTION: EXPRESSION TO THE EPIDERMIS OF TRANSGENIC ANIMALS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fulbright & Jaworski
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.A.
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/03985
; FILING DATE: 19930428
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.

```

THIS PAGE BLANK (USPTO)

```
/
/ REGISTRATION NUMBER: 32,714
/ REFERENCE/DOCKET NUMBER: D-5478
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 713/651-5325
/ TELEFAX: 713/651-5246
/ TELEX: 762829
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 24979 base pairs
/ TYPE: NUCLEIC ACID
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ PCT-US93-03985-3

alignment_scores:
  Quality: 80.50      Length: 72
  Ratio: 1.872      Gaps: 1
  Percent Similarity: 59.722      Percent Identity: 31.944

alignment_block:
US-09-327-750D-33 x PCT-US93-03985-3  ..
Align seg 1/1 to: PCT-US93-03985-3 from: 1 to: 24979

50 ArgGlyArgArgArgPheArgValArgGlnProIleSerHisTyrAr 66
||||| |||||:||||:||||: ||
5088 CGCGGAGCGCGCGGCGCTACGCTCAGGAGCTACGCGCGCGCGCAGCG 5137
66 qTrpAspLeuMethHisArgValGlyGluProGlnGlyArgMetArgGluG 83
|||| ||| :||| ||||: ||| :|||
5138 GTGGGGCTATGGCGCGCGGAGAGAGCTACCCGCGCGCGCGCGCAGGCG 5187
83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
:||||: ||||: ||| :|||: |||
5188 GGGAGTGGAGCGGTTATGGCAGCGGCTGCGCGCGCGGTGGCGGAGCTA 5237
100 ...ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPr 115
||||| |||||
5238 CGGAGGAGCGCGCAGAGCGCGCGCGGATCTCGCGCGTGCAGATCATCC 5287
115 oProHisHisAspHis 120
| |||:||||
5288 AGACCTCCACCAACAC 5303
```


248 AGAGAAACCGTAGCGGTTCCGCGTTAGGAGCCCATCTCTGAGTATAG 297
 66 gTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 298 ATGGGACATATGTCATAGGCTGGAGAGCCACAGGCAAGCATGAGAGAGG 347
 83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 348 AGAATATGGAAGGATTGGGAGAGGTGAGACAGCTGATGGAAAGCTG 397
 100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 398 AGGAAAAGCAGTTGACTCATAGTCTCGCGGCAGTCAGCACTGATCCCC 447
 116 oHisHisAspHisAspGluPheCysLeuMetPro 128
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 448 TCACCATGACCATCAGCATGAGTTTGCCTTATGCCC 484

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI58581

seq_documentation_block:

ID AAI58581 standard; cDNA; 862 BP.

XX AC AAI58581;

XX DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 784.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.

XX PR 25-APR-2000; 2000US-0552317.

XX PR 09-JUL-2000; 2000US-0598042.

XX PR 19-JUL-2000; 2000US-0620312.

XX PR 03-AUG-2000; 2000US-0653450.

XX PR 14-SEP-2000; 2000US-0662191.

XX PR 19-OCT-2000; 2000US-0693036.

XX PR 29-NOV-2000; 2000US-0727344.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren P, Wang D;

XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX PI Zhao QA, Zhou P, Goodrich R, Dmanac RT;

XX DR WPI; 2001-442253/47.

XX DR P-PSDB; AAM39425.

XX PT Novel nucleic acids and polypeptides, useful for treating disorders

XX PT such as central nervous system injuries -

XX PS Claim 1; SEQ ID NO 784; 10078pp; English.

XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and

XX CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,

XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful

XX CC in gene therapy. A composition containing a polypeptide or polynucleotide

XX CC of the invention may be used to treat diseases of the peripheral nervous

XX CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX SQ Sequence 862 BP; 235 A; 183 C; 238 G; 206 T; 0 other;

alignment_scores:

Quality: 468.00 Length: 129
 Ratio: 4.000 Gaps: 2
 Percent Similarity: 90.698 Percent Identity: 68.217

alignment_block:

US-09-327-750D-33 x AAI58581 ..

Align seg 1/1 to: AAI58581 from: 1 to: 862

1 MetGluSerLysAspGlnGlyAla...LysAsnLeuAsnMetGluAsnAs 16
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 226 ATGGAGTCCAAAGAGGAGCAGCGGTTAAACAATCTCATCTGCGGAAATGT 275
 16 pHisGlnLysLysGluGluLysGluLysProGlnAspThrIleLysA 33
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 276 CAACCAAGGAAATGATGAAAAAGATGAAAGGAGCAAGTTGCTAATAAAG 325
 33 rgGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 326 GGGAGCCCTTGGCCCTACCT...TTGAATGTTAGTGAATACTGTGTGCCT 372
 50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 373 AGAGGAAACCGTAGCGGTTCCGCGTTAGGAGCCCATCTCTGAGTATAG 422
 66 gTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 423 ATGGGACATATGTCATAGGCTGGAGAGCCACAGGCAAGCATGAGAGAGG 472
 83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 473 AGAATATGGAAGGATTGGGAGAGGTGAGACAGCTGATGGAAAGCTG 522
 100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 523 AGGAAAAGCAGTTGAGTTCATAGTTTGGCGGCAGTCAGCACTGATCCCC 572
 116 oHisHisAspHisAspGluPheCysLeuMetPro 128
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 573 TCACCATGACCATCAGCATGAGTTTGCCTTATGCCC 609

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF59611

seq_documentation_block:

ID AAF59611 standard; cDNA; 898 BP.

XX AC AAF59611;

XX DT 24-APR-2001 (first entry)

XX DE Human cell cycle and proliferation protein CCYPR-22 cDNA, SEQ ID NO:76.

XX KW Cell cycle and proliferation protein; CCYPR; human; agonist;

XX KW antagonist; gene therapy; detection; gene therapy;

XX KW transgenic animal disease model; immune disorder;

XX KW developmental disorder; cell signalling disorder;

XX KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;

XX KW arteriosclerosis; asthma; allergy; diabetes mellitus;

alignment_scores:
Quality: 466.50 Length: 129
Ratio: 4.092 Gaps: 3
Percent Similarity: 88.372 Percent Identity: 70.543

alignment_block:
US-09-327-750D-33 x AAH45143 ..
Align seg 1/1 to: AAH45143 from: 1 to: 792

1 MetGluSerLysAspGlnGlyAlaLysAsn...LeuAsnMetGluAsnAs 16
173 ATGGAGTCCCAAGAGAACTAGCTAAACAGTCTCAGCATGGGAAATGC 222
16 pHisGlnLysLysGluGluLysGluGluLysProGlnAspThrIleLysA 33
223 CAACCAAGAAATCAAGAAAGGAG.....CAAGTTCCTAATAAAG 263
33 rgGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
264 GGGAGCCCTTGGCCCTCCT...TTGGATGCTGTGAATCTGTGCCT 310
50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
311 AGAGAAATCGTAGCGGTTCCGCGTTAGCAGGCCCATCTGCAGTATAG 360
66 gTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83
361 ATGGATATGATGATAGCTTGGAGAACACAGGCAAGGATGAGAGAAG 410
83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
411 AGAATATGAAAGGATTTGGGAGGGGCTGAGACAGCTGATGGAAAGCTG 460
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
461 AGGAAAGAGGTTGAGTCTAGTCTCGCGGCACTGACGACTGACCCCC 510
116 oHisHisAspHisAspGluPheCysLeuMetPro 128
511 TCACCATGACCATCATGATGAGTTTGCCTTATGCC 547

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT: AAC03880

seq_documentation_block:
ID AAC03880 standard; cDNA; 662 BP.
XX AAC03880;
AC AAC03880;
XX 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 3878.
DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
KW Homo sapiens.
XX EP1033401-A2.
XX 06-SEP-2000.
XX 21-FEB-2000; 2000EP-0200610.
XX 26-FEB-1999; 99US-0122487.
XX (GEST) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI: 2000-500381/45.
XX P-PSDB: AAG03874.
XX

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
PS Claim 1: SEQ ID 3878; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
XX vectors.
SQ Sequence 662 BP; 177 A; 144 C; 192 G; 137 T; 12 Other;

alignment_scores:
Quality: 455.00 Length: 129
Ratio: 4.027 Gaps: 2
Percent Similarity: 87.597 Percent Identity: 66.667

alignment_block:
US-09-327-750D-33 x AAC03880 ..
Align seg 1/1 to: AAC03880 from: 1 to: 662

1 MetGluSerLysAspGlnGlyAla...LysAsnLeuAsnMetGluAsnAs 16
207 ATGGAGTCCCAAGAGAAAGAGGAGGCTTAAACAATCTCATCTGGGAAATGT 256
16 pHisGlnLysLysGluGluLysGluGluLysProGlnAspThrIleLysA 33
257 CAACCAAGAAATCAAGAAAGATGAAAGAGGCAAGTGCCTGAGTATAG 306
33 rgGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
307 GGGAGCCCTTGGCCCTACCT...TTGATGCTGTGTAATCTGTGCCT 353
50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
354 AGAGAAATCGTAGCGGTTCCGCGTTAGCAGGCCCATCTGCAGTATAG 403
66 gTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83
404 ATGGATATGATGATAGCTTGGAGAACACAGGCAAGGATGAGAGAAG 453
83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
454 AGAATATGAAAGGATTTGGGAGGAGGTGAGACAGCTGATGGAAAGCTG 503
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
504 AGGAAAGAGGTTGAGTCTAGTCTCGCGGCACTGACGACTGACCCCC 553
116 oHisHisAspHisAspGluPheCysLeuMetPro 128
554 TCACCATGACCATCATGATGAGTTTGCNNWATGCC 590

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT: AAC03879

seq_documentation_block:
ID AAC03879 standard; cDNA; 698 BP.
XX AAC03879;
XX AC AAC03879;
XX

06-OCT-2000 (first entry)

Human secreted protein 5' EST, SEQ ID NO: 3877.

Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation; gene therapy; chromosome mapping; ss.

Homo sapiens.

EP1033401-A2.

06-SEP-2000.

21-FEB-2000; 2000EP-0200610.

26-FEB-1999; 99US-0122487.

(GEST) GENSET.

Dumas Milne Edwards J, Duclert A, Giordano J;

WPI: 2000-500381/45.

P-PSDB; AAC03873.

New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures -

Claim 1; SEQ ID 3877; 7lpp + CD-ROM; English.

The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. An ORF has been identified within the sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.

Sequence 698 BP; 187 A; 151 C; 210 G; 139 T; 11 other;

alignment_scores:
Quality: 455.00 Length: 129
Ratio: 4.027 Gaps: 2
Percent Similarity: 87.597 Percent Identity: 66.667

alignment_block:
US-09-327-750D-33 x AAC03879 ..

Align seg 1/1 to: AAC03879 from: 1 to: 698

1 MetGluSerLysAspGlnGlyAla...LysAsnLeuAsnMetGluAsnAs 16
|||||..... |||| :||| :||| :||| :||| :||| :|||
243 ATGGAGTCCAAAGAGGACGCGTTAAACAATCATCGTGGAATAATGT 292

16 pHisGlnLysLysGluLysGluLysProGlnAspThrIleLysA 33
:|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
293 CAACGACGAAATATGTAAGAAGATGAAGAAGCAGCAAGTCGTAATAAG 342

33 rgGluProValValAlaPheThrPheGluAlaGlyLysAsnCysAlaPro 49
:|||||:|||||:|||||:|||||:|||||:|||||:|||||
343 GGGAGCCCTTGGCCCTACCT...TTGRATGYTRGTGAATACTGTGCCT 389

50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyAr 66
:|||||:|||||:|||||:|||||:|||||:|||||:|||||
390 AGAGAAATCGTAGCGGTTCCGCTTAGCAGGCCATCCCTCGACTATAG 439

[illegible]

Sequence 700 BP; 177 A; 188 C; 203 G; 132 T; 0 other;

Quality:	173.00
Ratio:	2.190

THIS PAGE BLANK (USPTO)

```
alignment_block:
US-09-327-750D-33 x AAF21748  ..
Align seg 1/1 to: AAF21748 from: 1 to: 917

18 GlnLysGluGluLysGluLysProGlnAspThrIleLysArgG1 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
306 CAGGAAACCAAGAGATGGAGCGCTTATCGACAATGAGAGAGACCG 355
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 uProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
356 CCCTTTG.....GGAGGAGGTGAAGGCCACCGCCT...G 387

51 lYgLYArgArgPheArgValArgGlnProIleSerHisTyrArgTIP 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
388 CAGGAATCGACGGGACAGCGCTCGCGACTTCCCTTATTTTCGATGG 437
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
68 AspLeuMethHisArg...ValGlyGluProGlnGlyArgMetArgGluG1 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
438 GCCATACCCCAATAGGCAGATCAATGATGGGATGGGTGA...GATGAGA 484
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
83 uAsnValGlnArgPheGlyClyAspMetArgGlnLeuMetGluLysLeuA 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
485 TGATATGGAATATTCATGAGGAGATGAGAGAAATCAGAGAAACTTA 534
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
100 rGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAsp...Pro 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
535 GGAGCTGCAGTTGAGGAATGTGCTGCTATCCTTATGGGGGAGCTCTCT 584
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 ProHisHisAspHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
585 AATCACCACCATCATCATGATGATTTTCCTTATGCCT 623
```

seq_name: /SID32/gcgdata/geneseq/geneseq/NA2000.DAT:AAC01005

```
seq_documentation_block:
ID AAC01005 standard; cDNA; 532 BP.
AC AAC01005;
DT 06-OCT-2000 (first entry)
DE Human secreted protein 5' EST, SEQ ID NO: 1003.
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX EP1033401-A2.
XX 06-SEP-2000.
XX 21-FEB-2000; 2000EP-0200610.
XX 26-FEB-1999; 99US-0122487.
XX (GEST ) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI: 2000-500381/45.
XX P-PSDB: AAG00999.
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 1; SEQ ID 1003; 71pp + CD-ROM; English.
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified within the
```

CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 532 BP; 151 A; 121 C; 163 G; 91 T; 6 other;

```
alignment_scores:
Quality: 135.50 Length: 120
Ratio: 1.831 Gaps: 5
Percent Similarity: 61.667 Percent Identity: 32.500

alignment_block:
US-09-327-750D-33 x AAC01005  ..
Align seg 1/1 to: AAC01005 from: 1 to: 532

4 LysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHisGlnLysLy 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
192 AACAAACCCAGAAAAAAATCTCATCATGGCAAAATATTCACCCAGGAAA 241
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20 sGluGluLysGluLysProGlnAspThrIleLysArgGluProValV 37
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
242 CGAAGAGATGGACGACGCTATGCAGATGAGAGAGAGACGCCCTTTG. 290
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
37 alaProThrPheGluAlaGlyLysAsnCysAlaProArgGlyGlyArg 53
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
291 .....GGAGGAGGTGAAGGCCACCGCCT...GCAGGAAAT 323
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
54 ArgArgPheArgValArgGlnProIleSerHisTyrArgTrpAspLeu 70
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
324 CGACGGGAMAGCTCGCGAYTTGCCCTTAATTTTCGATGGGCCATACC 373
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
70 tHisArg...ValGlyGluProGlnGlyArgMetArgGluGluAsnVal 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
374 CAATAGGCAGATCAATGATGGGATGGT...GRAGATGGAGATGATATGG 420
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
86 lnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArgGluArg 102
:: ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:
421 AAATATTCATGGAGGAGATGAGAGAAATCAGAGAAACCTTAGGGAGCTG 470
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
103 GlnLeuSerHisSerLeuArg...AlaValSerThrAspProHisH1 118
||| ::| |||||:|||||:|||||:|||||:|||||:|||||:|||||:
471 CAKTTGAGGAATGTCTGCGTATCTCTATGGGGGAKCTCTCTAATACCA 520
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
118 sAspHisHis 121
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
521 TGACCATCAT 530
```


/note="retinoic acid reduced gene expression"

/codon_start=1
/product="REX-3"
/protein_id="AAC61929.1"
/db_xref="GI:3510643"
/translation="MENDHQKKEEKEKPQDITRREPAVALISEAGKNCAPRGGRRRF
RYQPIAHYRNDLMQVRGEPQGRMEENVQRFQDLMEKLRERQLSHSLRAVSTD
PPHDDHDFCLMP"
BASE COUNT 238 a 181 c 251 g 168 t
ORIGIN

alignment_scores:
Quality: 640.00 Length: 128
Ratio: 5.161 Gaps: 0
Percent Similarity: 96.875 Percent Identity: 92.188

alignment_block:

US-09-327-750D-33 x AF051347

Align seg 1/1 to: AF051347 from: 1 to: 838

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHi 17
|||||
193 ATGGAGTCCAAAGATCAAGCGCTGAAATAATCTCAACATGGAGAATGACCA 242
17 sGlnLysLysGluGluLysGluGluLysProGlnAspThrIleLysArg 34
|||||
243 TCAGAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 292
34 luProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
|||||
293 AGCCAGCTGTGGCCCTCATCTCCGAGCTGGCAAAACTGCGGCCCTAGA 342
51 GlyGlyArgArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67
|||||
343 GGAGGTCGCAGCGGTTCCGGGTTCCGGAGCCCATCGCTCACTATAGATG 392
67 pAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluGluA 84
|||||
393 GGACCTGATCGAGAGGGTTGGGAGCCGCCAGGAGGAGGATGAGAGGAGA 442
84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
|||||
443 ACGTACAGAGGTTTGGGGGTGATGTGAGACAGCTCATGGAGAGCTCAGG 492
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||
493 GAAAGGCAGCTGAGCCACAGCTCGCGGCGGTTAGCAGTACCCGCGCTCA 542
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
543 TCATGACCACCATGATGAGTTTGCCTCATGCCCC 576

seq_name: gb_ro:AF097439

seq_documentation_block:

LOCUS AF097439 785 bp mRNA ROD 13-APR-1999
DEFINITION Mus musculus brain expressed x-linked protein 2 (Bex2) mRNA,
complete cds.

ACCESSION AF097439
VERSION AF097439
KEYWORDS AF097439.1 GI:4580591

SOURCE house mouse.

ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 785)

AUTHORS Brown,A.L. and Kay,G.F.

TITLE Bex2, a gene with increased expression in parthenogenetic embryos,

is a member of a novel gene family on the mouse X chromosome

JOURNAL Hum. Mol. Genet. 8 (4), 611-619 (1999)

MEDLINE 99172070

REMARK

REFERENCE 2 (bases 1 to 785)

AUTHORS Brown,A.L. and Kay,G.F.

TITLE Direct Submission

JOURNAL Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia

FEATURES Location/Qualifiers

Source

1..785

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/db_xref="dbEST:W48832"

/chromosome="X"

/map="near plp"

/dev_stage="embryo; 15.5 dpc"

1..785

/gene="Bex2"

139..528

/gene="Bex2"

/codon_start=1

/product="brain expressed x-linked protein 2"

/protein_id="AAD24430.1"

/db_xref="GI:4580592"

/translation="MESKVEQGVKNLNMNDHQEKEEKEEKPODASKRDPVIALPFEA
GDYVYRGRRRRFRVQPIVHYRNDLMHRVGEPOGRMEENVQRFQDLMEKLRERQLSHSLRAVSTD
PPHDDHDFCLMP"
BASE COUNT 212 a 163 c 235 g 175 t
ORIGIN

alignment_scores:

Quality: 601.50 Length: 129

Ratio: 4.971 Gaps: 1

Percent Similarity: 93.798 Percent Identity: 86.822

alignment_block:

US-09-327-750D-33 x AF097439

Align seg 1/1 to: AF097439 from: 1 to: 785

1 MetGluSerLys...AspGlnGlyAlaLysAsnLeuAsnMetGluAsnAs 16

139 ATGGAGTCCAAAGTGGAAACAAAGCGTCAAAATCTCAACATGGAGAATGA 188

16 pHisGlnLysLysGluGluLysGluGluLysProGlnAspThrIleLysA 33

189 CCATCAGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 238

33 rgGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49

239 GGGATCCGATTGTGGCCCTGCCTTTCGAAGCTGGAGACTACTACGTCCT 288

50 ArgGlyGlyArgArgPheArgValArgValArgGlnProIleSerHisTyrAr 66

289 AGAGGAGTCCGAGCGGTTCCGGGTTCCGAGCCCATCTGTCACACTACAG 338

66 gTrpAspLeuMethHisArgValGlyGluProGlnGlnGlyArgMetArgGlu 83

339 ATGGGACCTGATGATAGGGTTGGGAGCCGCCAGGAGGAGGAGGAGGAGG 388

83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99

389 AGACGTACAGAGGTTGGGAGTATGTGACAGCTCATGGAGAAAGCTG 438

100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116

439 AGGAAAGGAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 488

116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128

489 TCATCATGACCACCATGATGAGTTTGCCTTATGCC 525

Erratum: [[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]

seq_name: gb_ro:BC003254

seq_documentation_block:

LOCUS BC003254 2901 bp mRNA ROD 20-FEB-2001

DEFINITION Mus musculus, Similar to dentatorubral pallidolusian atrophy, clone MGC:5758, mRNA, complete cds.

ACCESSION BC003254

VERSION BC003254.1 GI:13096906

KEYWORDS MGC.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS 1 (bases 1 to 2901)

TITLE Direct Submission

JOURNAL Strausberg R.

REMARK Submitted (20-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Jeffrey Green M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: villalon@bcm.tmc.edu

Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK plate: 9 Row: 0 Column: 6.

FEATURES

source

1..2901

Location/Qualifiers

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="MGC:5758 IMAGE:3500522"

/tissue_type="mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."

/clone_lib="NCI_CGAP_Mam6"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

755..2359

/codon_start=1

/product="Similar to dentatorubral pallidolusian atrophy"

/protein_id="AAH03254.1"

/db_xref="GI:13096907"

/translation="MVQFLLPLPLKPPHVVHTPLHPRAKHEPTPSHQSLQSP PQLPFPLSSPPAAPTGTPLTATQIKOEPAEYEPSPVPARSPSPKVVDPVS HASQARENKHLDRCFNSCSDLYFPLEGSKLAKRADLYEKVRRCAEQAREKE RREARENKREERERERSKLVKLAQGRAPVECPGLVPVHPRPPEGSAAVTPP YLGDPDTALRUSEYARHVMSPGNRNHPFVYGLGADPGLLYNVPALYSSDPAARE RREAREDRDLKPGFEVKPSELEPLHVGPGGLDPFRHGLGLAQGPGLHPFP FPLSGPLERLALAAGPALRPMDSYAEALRAERQAERVAALGNLPLARLQMLNVT PHHQSHLHSHLHQQDAIHAASVSHPLDIDPLASGSHLTRIPYPAGTLPNLLPH PLHNEVLRHOLFAPYRDLPASLAPMSAAHQLQAMHAQSAELORLAEQQOHLHAH HPLHSVPLPAQEDYSHLKESDKPL"

BASE COUNT 623 a 958 c 791 g 529 t

ORIGIN

alignment_scores:

Quality: 582.00 Length: 118

Ratio: 5.061 Gaps: 0

Percent Similarity: 97.458 Percent Identity: 92.373

alignment_block:

US-09-327-750D-33 x BC003254

Align seg 1/1 to: BC003254 from: 1 to: 2901

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspH1 17

|||||

223 ATGGAGTCCAAAGATCAAGGGGTGAAATACTCAACATGGAGATGACCA 272

17 sGlnLysGluGluLysGluGluLysProGlnAspThrIleLysArgG 34

|||||

273 TCAGAAAAAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 322

34 luProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50

|||||

323 AGCCAGCTGTGGCCCTGACCTCCGAGGCTGGCAAAAACCTGTGCACCTAGA 372

51 GlyGlyArgArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67

|||||

373 GGAGTCCGAGGCGGTTCGGGTTCCGGCTCGGCAGCCCATCGCTCACTATAGTG 422

67 pasLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluGluA 84

|||||

423 GGACCTGATGCAGAGGTTGGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 472

84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100

|||||

473 ACGTACAGAGGTTGGGGGTTGAGTGTGAGCAGCTCATGGAGAGCTGAGG 522

101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH1 117

|||||

523 GAAAGGAGCTGAGCCACACCTTGGGGCGGTAGCACTGACCCGCTCA 572

117 shis 118

|||||

573 TCAT 576

seq_name: gb_pr:AF183416

seq_documentation_block:

LOCUS AF183416 642 bp mRNA PRI 02-SEP-2000

DEFINITION Homo sapiens ovarian granulosa cell 13.0 kDa protein hGR74 homolog mRNA, complete cds.

ACCESSION AF183416

VERSION AF183416.1 GI:9963770

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 642)

AUTHORS Peng, Y., Qian, B., Tu, Y., Xu, S., Han, Z., Fu, G. and Chen, Z.

TITLE A novel gene expressed in human adrenal gland

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 642)

AUTHORS Peng, Y., Gu, W., Huang, C., Xu, S., Han, Z., Fu, G. and Chen, Z.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-1999) Chinese National Human Genome Center at Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China

FEATURES

source

1..642

/organism="Homo sapiens"

/db_xref="taxon:9606"

/tissue_type="adrenal gland"

3..380

/note="HGR74-h protein"

/codon_start=1

/evidence=not_experimental

/product="ovarian granulosa cell 13.0 kDa protein hGR74 homolog"

/protein_id="AA09685.1"

/db_xref="GI:9963771"

TITLE
JOURNAL

Direct Submission
Submitted (21-JAN-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk

COMMENT

On May 13, 1999 this sequence version replaced gi:4582117.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.

This sequence was generated from part of bacterial clone contigs of
human chromosome X, constructed by the Sanger Centre Chromosome X
Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/ChrX>

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RPI-198P4 is
from the library RPCI-1 constructed at the Roswell Park Cancer
Institute by the group of Pieter de Jong. For further details see
<http://bacpac.med.buffalo.edu/>

VECTOR: pCYPAC2

IMPORTANT: This sequence is not the entire insert of clone
RPI-198P4. It may be shorter because we only sequence overlapping
sections once, or longer because we arrange for a small overlap
between neighbouring submissions.
The true left end of clone LLOXNC01-221P2 is at 35615 in this
sequence. The true right end of clone LLOXNC01-73E8 is at 100 in
this sequence.

FEATURES

source

Location/Qualifiers

1..35714
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q22"
/clone="RPI-198P4"
/clone.lib="RPCI-1"

repeat_region
1..258
/note="L1P3 repeat: matches 5666. .5925 of consensus"
256..537
/note="L1P3 repeat: matches 5391. .5672 of consensus"
1491..1749
/note="AluSc repeat: matches 39. .297 of consensus"
2182..3810
/note="L1P3 repeat: matches 4533. .6150 of consensus"
5169..5687
/note="L1MA2 repeat: matches 4798. .6308 of consensus"
6976..7665
/note="L1MA2 repeat: matches 4113. .4798 of consensus"
8039..9213
/note="L1MA2 repeat: matches 2945. .4113 of consensus"
9523..10437
/note="L1MA2 repeat: matches 2016. .2945 of consensus"
10736..11494
/note="L1MA2 repeat: matches 1262. .2016 of consensus"
11807..12251
/note="L1MA2 repeat: matches 310. .755 of consensus"
12259..12331
/note="HERVFB21 repeat: matches 5. .77 of consensus"
12332..12796
/note="L1TR21A repeat: matches 1. .490 of consensus"
16162..16371
/note="105 copies 2 mer tt 55% conserved"
17121..17424
/note="MER33 repeat: matches 1. .322 of consensus"

repeat_region
18259..18290
/note="16 copies 2 mer ca 87% conserved"
18505..18672
/note="MER63A repeat: matches 1. .181 of consensus"
19478..19814
/note="MER2 repeat: matches 1. .345 of consensus"
21061..21101
/note="HERVFB21 repeat: matches 42. .85 of consensus"
21247..21781
/note="MLT1-INTERNAL repeat: matches 44. .568 of consensus"
24465..24492
/note="Single clone region"
Single clone region"
24531..24562
/note="Weak data"
Weak data"
24563
/note="Single clone region"
Single clone region"
24711..24865
/note="region covered by subclones from a PAC DNA PCR
product only at 12x coverage."
region covered by subclones from a PAC DNA PCR product only
at 12x coverage."
24758..25460
/note="CpG island"
/evidence="not_experimental"
24875..24940
/note="33 copies 2 mer cc 66% conserved"
26795..26926
/note="66 copies 2 mer ct 63% conserved"
27088..27133
/note="23 copies 2 mer tg 100% conserved"
27461..27777
/note="L1MB8 repeat: matches 5851. .6171 of consensus"
27802..28232
/note="L1MC/D repeat: matches 5192. .5403 of consensus"
28536..28930
/note="L1MC/D repeat: matches 4802. .5192 of consensus"
29240..29341
/note="TIGGER1 repeat: matches 484. .587 of consensus"
29603..30740
/note="TIGGER1 repeat: matches 587. .1757 of consensus"
31027..31365
/note="TIGGER1 repeat: matches 1757. .2105 of consensus"
31664..31691
/note="TIGGER1 repeat: matches 2105. .2132 of consensus"
31809..32015
/note="TIGGER1 repeat: matches 2132. .2246 of consensus"
32323..32389
/note="TIGGER1 repeat: matches 2246. .2313 of consensus"
34353..34636
/note="MLT1-INTERNAL repeat: matches 607. .889 of
consensus"
35149..35376
/note="MLT1-INTERNAL repeat: matches 1001. .1233 of
consensus"
35523..35670
/note="MLT1C repeat: matches 8. .159 of consensus"
BASE COUNT 9708 a 7094 c 6805 g 12107 t
ORIGIN

alignment_scores:

Quality: 471.50 Length: 129
Ratio: 4.100 Gaps: 3
Percent Similarity: 89.147 Percent Identity: 70.543

alignment_block:

US-09-327-750D-33 x HS198P4/rev ..

Align seg 1/1 to reverse of: HS198P4 from: 1 to: 35714

```

1 MetGluSerLysAspGlnGlyAlaLysAsn...LeuAsnMetGluAsnAs 16
24358 ATGGAGTCCAAAGAGAACGACGAGTAAACAGCTCTCAGCATGGAAAATGC 24309
16 pHisGlnLysLysGluGluLysGluLysProGlnAspThrIleLysA 33
24308 CAACCAAGAAATGAAGAAAGGAG.....CAAGTTGCTAATAAAG 24268
33 rGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
24267 GGGAGCCCTTGGCCCTCCCT...TTGGATGCTGGTCACTACTGTGTGCCT 24221
50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
24220 AGAGGAATCGTAGGCGGTTCGCGGTAGGACGCCCATCTGCAGTATAG 24171
66 gtrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGlu 83
24170 ATGGGATATGATGATGATGCTTGGAGAACCCACAGGCAAGGATGAGAGAAG 24121
83 LuAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
24120 AGAATATGAAGAGATTGGGAGGAGTGCAGACGCTGATGGAAAGCTG 24071
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
24070 AGGGAAGAGCTGAGTGCATAGTCTGCGGCGAGTCAGCACTGACCCGCC 24021
116 OHHisAspHisAspGluPheCysLeuMetPro 128
24020 TCACCATGACCATCATGATGATGAGTTTGCCTTATGCC 23984
seq_name: gb_pr:AF251053

```

```

seq_documentation_block: 744 bp mRNA PRI 15-APR-2001
LOCUS AF251053
DEFINITION Homo sapiens X-linked protein mRNA, complete cds.
ACCESSION AF251053
VERSION AF251053.1 GI:13625167
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Mao, Y., Xie, Y., Zhou, Z., Zhao, W., Wang, W., Huang, Y.,
Wang, S., Tang, R., Chen, X. and Wu, C.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2000) Institute of Genetics, School of Life
Sciences, Fudan University, 220 Handan Road, Shanghai 200433, P.R.
China
FEATURES
source
1..744
/organism="Homo sapiens"
/db_xref="taxon:9606"
101..487
CDS
101..487
/codon_start=1
/product="X-linked protein"
/protein_id="AAK34943.1"
/db_xref="GI:13625168"
/translation="MESKEERALNNLIVENNQENDEKDEQVANKGEPLALPLNYS
EYCVPRGNRRFRVQPIQYRWDMHRLGEPQARMREENMERIGEVQLMEKLREK
QLSHSLRAVSTDPPHHDHDFCLMP"
BASE COUNT 222 a 134 c 208 g 180 t
ORIGIN
alignment_scores:
Quality: 468.00 Length: 129
Ratio: 4.000 Gaps: 2
Percent Similarity: 90.698 Percent Identity: 68.217
alignment_block:
US-09-327-750D-33 x AX078272
Align seg 1/1 to: AX078272 from: 1 to: 898

```

```

US-09-327-750D-33 x AF251053
Align seg 1/1 to: AF251053 from: 1 to: 744
1 MetGluSerLysAspGlnGlyAla...LysAsnLeuAsnMetGluAsnAs 16
101 ATGGAGTCCAAAGAGAACGACGCGTTAAACAATCTCATCGTGGAAAATGT 150
16 pHisGlnLysLysGluGluLysGluLysProGlnAspThrIleLysA 33
151 CAACCAAGAAATGATGAAAAGATGAAAAGGCAAGTTCCTAATAAAG 200
33 rGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
201 GGGAGCCCTTGGCCCTTACCT...TTGAATGTTAGTGAATCTGTGTGCCT 247
50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyrAr 66
248 AGAGGAACCCGTAGGCGGTTCGCGGTAGGACGCCCATCTGCAGTATAG 297
66 gtrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGlu 83
298 ATGGACATGAATGCATAGCTTGGAGAGCCACAGGCAAGATGAGAGAG 347
83 LuAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
348 AGAATATGAAGAGATTGGGAGGAGGTGAGACGCTGATGGAAGAGCTG 397
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
398 AGGGAAGAGCTGAGTGCATAGTCTGCGGCGAGTCAGCACTGATCCCC 447
116 OHHisAspHisAspGluPheCysLeuMetPro 128
448 TCACCATGACCATCAGTATGAGTTTGCCTTATGCC 484
seq_name: gb_pat:AX078272
seq_documentation_block: 898 bp DNA PAT 22-FEB-2001
LOCUS AX078272
DEFINITION Sequence 76 from Patent WO0107471.
ACCESSION AX078272
VERSION AX078272.1 GI:13157963
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Hui man, J. L., Lal, P., Tang, Y. T., Yue, H., Au-Young, J., Bandman, O.,
Azinval, Y., Yang, J., Lu, D. A., Baughn, M. R., Patterson, C. and Shah, P.
TITLE Cell cycle and proliferation proteins
JOURNAL Patent: WO 0107471-A 76 01-FEB-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
source
1..898
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 5664154CB1"
BASE COUNT 250 a 186 c 251 g 211 t
ORIGIN
alignment_scores:
Quality: 468.00 Length: 129
Ratio: 4.000 Gaps: 2
Percent Similarity: 90.698 Percent Identity: 68.217
alignment_block:
US-09-327-750D-33 x AX078272
Align seg 1/1 to: AX078272 from: 1 to: 898

```

```

1 MetGluSerLysAspGlnGlyAla...LysAsnLeuAsnMetGluAsnAs 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
254 ATGAGTCCAAAGGAAGCAGCTTAAACAATCTCATCTGTTGAAATGT 303

16 pHisGlnLysGluGlyLysGluGlyLysProGlnAspThrIleLysA 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
304 CAACGAGAAATGATGAAAAGATGAAAAGGAGCAAGTTCTTAATAAAG 353

33 rgGluProValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
354 GGGAGCCCTGGCCCTACCT...TTGATGTTAGTGAATACTGTGTCCT 400

50 ArgGlyArgArgArgPheArgValArgGlnProIleSerHisTyAr 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
401 AGAGAAACCGTAGCGGTTCCGCGTTAGCAGCCCATCTCGCAGTATAG 450

66 gTriPAspLeuMethHisArgValGlyGluProGlnGlyArgMetArgGlu 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
451 ATGGGACATAATGCATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGAG 500

83 luAsnValGlnArgPheGlyCysAspMetArgGlnLeuMetGluLysLeu 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
501 AGAATATGGAAGGATTTGGGAGGAGGTGAGACAGCTGATGGAAGAGCTG 550

100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
551 AGGCAAAAGCAGTTGACTCATAGTCTCGGCGCAGTCAGCACTGATCCCC 600

116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
601 TCACCATGACCATCAGCATGAGTTTGCCTTATGCC 637

```

seq_name: gb_pr:HSV870H8

```

seq_documentation_block: 31321 bp DNA PRI 23-NOV-1999
LOCUS HSV870H8
DEFINITION Human DNA sequence from cosmid V870H8, between markers DXS366 and
DXS87 on chromosome X contains ESTs.
ACCESSION 270233
VERSION 270233.1 GI:1235542
KEYWORDS X.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 31321)
AUTHORS Whiteley,M.
TITLE Direct Submission
JOURNAL Submitted (19-MAR-1995) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, UK. E-mail enquires: humquery@sanger.ac.uk
COMMENT IMPORTANT: This sequence is not the entire insert of clone V870H8.
It may be shorter because we arrange for a small overlap between
once, or longer because we
neighbouring submissions.
The true left end of clone V870H8 is at 1 in this sequence. The
true left end of clone V693A8 is at 3121.
V870H8 is from the human chromosome X-specific cosmid library.
FEATURES
source
1..31321
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="X"
/clone="CHC-870H8"
/clone_lib="SCCV"
475..701
/repeat_region
802..988
/repeat_region
2533..4138
/misc_feature
4189..4246
/repeat_region
/note="29 copies of 2 mer 91 % conserved"

```

```

repeat_region 7496..7874
/note="L1 element fragment"
repeat_region 8022..8108
/note="L1 element fragment"
repeat_region 8302..8493
/note="L1 element fragment"
repeat_region 8585..8671
/note="L1 element fragment"
repeat_region 8769..9010
/note="L1 element fragment"
repeat_region 9259..9384
/note="L1 element fragment"
repeat_region 9624..9731
/note="MLTIC element fragment"
repeat_region 9781..9948
/note="MLTIC element fragment"
repeat_region 9890..9948
/note="MLTIB element fragment"
repeat_region 10418..10592
/partial
/note="Alu repeat: matches 298..108 of consensus"
10593..10688
/partial
/note="Alu repeat: matches 96..1 of consensus"
10698..10781
/note="L1 element fragment"
repeat_region 11112..11403
/note="Alu repeat: matches 1..308 of consensus"
11922..12004
/note="L1 element fragment"
repeat_region 12062..12308
/note="L1 element fragment"
repeat_region 12423..12707
/partial
/note="Alu repeat: matches 302..1 of consensus"
12710..12916
/note="L1 element fragment"
repeat_region 12979..13081
/note="L1 element fragment"
repeat_region 13473..13532
/note="MLT2A1 element fragment"
repeat_region 15871..15928
/note="L1 element fragment"
repeat_region 18180..18417
/note="L1 element fragment"
repeat_region 18778..19056
/note="Alu repeat: matches 1..308 of consensus"
22279..22350
/note="L1 element fragment"
repeat_region 22449..22995
/note="L1 element fragment"
repeat_region 23239..23374
/note="L1 element fragment"
repeat_region 23371..23606
/note="MLTIB element fragment"
repeat_region 23395..23541
/note="MLT1A element fragment"
repeat_region 23705..23778
/note="MLT1A element fragment"
repeat_region 23805..23960
/note="L1 element fragment"
repeat_region 24011..24541
/note="L1 element fragment"
repeat_region 24638..24770
/partial
/note="Alu repeat: matches 150..1 of consensus"
24829..30975
/note="L1 element fragment"
BASE COUNT 9462 a 6147 c 6129 g 9583 t
ORIGIN
alignment_scores:

```

Quality: 468.00 Length: 129
Ratio: 4.000 Gaps: 2
Percent Similarity: 90.698 Percent Identity: 68.217

alignment_block:
US-09-327-750D-33 x HSV870H8 ..

Align seg 1/1 to: HSV870H8 from: 1 to: 31321

1 MetGluSerLysAspGlnGlyAla...LysAsnLeuAsnMetGluAsnAs 16
|||||.....:|||||.....:|||||

3508 ATGGAGTCAAGAGAGACGAGCGTTAAACAATCTCATCTGTGGAATGT 3557
|||||.....:|||||.....:|||||

16 pHisGlnLysLysGluGluLysGluLysProGlnAspThrIleLysA 33
:|||||.....:|||||.....:|||||

3558 CAACAGGAAATGATGAAAAAGATCAAAAGGACCAAGTGTCTAATAAG 3607
:|||||.....:|||||.....:|||||

33 rGluProValValAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
|||||.....:|||||.....:|||||

3608 GGGAGCCCTTGGCCCTTACCT...TTGAATGTTAGTGAATACTGTGTCCT 3654
|||||.....:|||||.....:|||||

50 ArgGlyGlyArgArgPheArgValArgGlnProIleSerHisTyraR 66
|||||.....:|||||.....:|||||

3655 AGAGGAACCGTAGGCGGTTCCCGTTAGCGACCCCATCTGCAGTATAG 3704
|||||.....:|||||.....:|||||

66 gTTPspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83
|||||.....:|||||.....:|||||

3705 ATGGGACATAATGCATAGCTTGGAGAGCCACAGGCAAGGATGAGAGAG 3754
|||||.....:|||||.....:|||||

83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
|||||.....:|||||.....:|||||

3755 AGAATATGGAAGAGATTTGGGAGGAGGTGAGACAGCTGATGGAAAAAGCTG 3804
|||||.....:|||||.....:|||||

100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||.....:|||||.....:|||||

3805 AGGGAAGACAGTGTAGTCATAGTCTGCGGCGAGTCAGCACTGATCCCC 3854
|||||.....:|||||.....:|||||

116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
|||||.....:|||||.....:|||||

3855 TCACCATGACCATCAGCATGAGTTTGCCTTATGCCC 3891

seq_name: gb_pr:AL133348

seq_documentation_block:
LOCUS AL133348 40584 bp DNA PRI 06-MAR-2000
DEFINITION Human DNA sequence from clone RPI-79p11 on chromosome Xq21.32-22.1.
Contains the gene for a novel protein similar to mouse Bex2
(brain-expressed X-linked protein 2), ESRs, STSS, GSSs and a
putative CpG island, complete sequence.

ACCESSION

AL133348

VERSION

AL133348.8 GI:7076386

KEYWORDS

HTG; Bex2; CpG Island.

SOURCE

human

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 40584)

REFERENCE

1 (bases 1 to 40584)

AUTHORS

Wilson,S.

TITLE

Direct Submission

JOURNAL

Submitted (02-MAR-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Feb 24, 2000 this sequence version replaced gi:6997869.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the

sequence is ambiguous, there is an annotation using the 'unsure'
feature key.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at

http://www.sanger.ac.uk/HGP/ChrX
RPI-79p11 is from the library RPI-1 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: PCYPAC2

IMPORTANT: This sequence is not the entire insert of clone
RPI-79p11. It may be shorter because we only sequence overlapping
sections once, or longer because we arrange for a small overlap
between neighbouring submissions.

The true left end of clone LLOXNC01-105G4 is at 40485 in this
sequence. The true right end of clone LLOXNC01-177E8 is at 100 in
this sequence.

FEATURES

source

Location/Qualifiers

1..40584

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="X"

/map="q21.32-22.1"

/clone="RPI-79p11"

/clone_lib="RPI-1"

19..292

/note="AluJo repeat: matches 3..275 of consensus"

470..1284

/note="L1MB5 repeat: matches 5318..6157 of consensus"

1295..1898

/note="L1PA15 repeat: matches 5556..6157 of consensus"

1899..2345

/note="L1PA15 repeat: matches 4922..5370 of consensus"

2323..2698

/note="L1MB5 repeat: matches 4960..5343 of consensus"

2927..3065

/note="MER74A repeat: matches 164..305 of consensus"

3655..9795

/note="L1PA4 repeat: matches 2..6142 of consensus"

9866..9982

/note="FLAM_C repeat: matches 6..124 of consensus"

9989..10254

/note="L1 repeat: matches 4027..4293 of consensus"

10253..10852

/note="L1M4 repeat: matches 4755..5333 of consensus"

10853..11260

/note="MLT1B repeat: matches 3..390 of consensus"

11261..11474

/note="L1M4 repeat: matches 4538..4755 of consensus"

11480..11563

/note="MSTB repeat: matches 1..80 of consensus"

11564..12475

/note="L1 repeat: matches 3039..3971 of consensus"

14463..14861

/note="match: GSS: Em:AQ729215"

15578..15854

/note="AluJo repeat: matches 6..289 of consensus"

complement(16007..16556)

/note="match: GSS: Em:AQ72810"

16197..16569

/note="L1M4 repeat: matches 4267..4627 of consensus"

16629..16714

/note="L1M4 repeat: matches 1578..1666 of consensus"

16902..17133

/note="L1M4 repeat: matches 0..242 of consensus"

17536..17728

/note="MIR repeat: matches 1..191 of consensus"

17746..18323

```

/notes="L2 repeat: matches 1599. .2175 of consensus"
18706. .18755
/notes="LIMB3 repeat: matches 6034. .6083 of consensus"
18756. .19236
/notes="MER66B repeat: matches 1. .486 of consensus"
19237. .21108
/notes="MER66-internal repeat: matches 4808. .6676 of
consensus"
21109. .21163
/notes="MT2B repeat: matches 391. .445 of consensus"
21164. .21555
/notes="MER66B repeat: matches 3. .391 of consensus"
21557. .21662
/notes="LIMB3 repeat: matches 6080. .6185 of consensus"
21665. .21761
/notes="LIM repeat: matches 5299. .5392 of consensus"
21762. .21929
/notes="LIMB8 repeat: matches 5902. .6063 of consensus"
21930. .22212
/notes="AluSg repeat: matches 3. .284 of consensus"
22213. .22321
/notes="LIMB8 repeat: matches 6063. .6173 of consensus"
22333. .22740
/notes="LIMD2 repeat: matches 5867. .6288 of consensus"
22744. .23094
/notes="LIMD repeat: matches 2. .342 of consensus"
23218. .23524
/notes="AluSg repeat: matches 1. .305 of consensus"
23603. .23947
/notes="LIM4 repeat: matches 4916. .5184 of consensus"
23948. .24218
/notes="AluJo repeat: matches 1. .282 of consensus"
24219. .24510
/notes="LIM4 repeat: matches 4601. .4916 of consensus"
24508. .24621
/notes="LIM4 repeat: matches 4374. .4488 of consensus"
24650. .25077
/notes="LIM4 repeat: matches 1. .466 of consensus"
25078. .25914
/notes="LIM4 repeat: matches 3497. .4346 of consensus"
25920. .26741
/notes="LIM4 repeat: matches 2202. .3065 of consensus"
26762. .26906
/notes="LIP7 repeat: matches 6001. .6145 of consensus"
26907. .27135
/notes="LIP5 repeat: matches 5674. .5902 of consensus"
27266. .27509
/notes="LIM4 repeat: matches 2164. .2080 of consensus"
27557. .27769
/notes="LIMEC repeat: matches 253. .460 of consensus"
28120. .28167
/notes="L2 repeat: matches 2703. .2750 of consensus"
29043. .29349
/notes="MER58B repeat: matches 1. .341 of consensus"
30389. .30452
/notes="16 copies 4 mer tata 89% conserved"
complement(30502)
/gene="dj79p11.1"
complement(join(30502. .31137,31514. .31589,32016. .>32134))
/notes="match: ESTs: Em:AA317587 Em:AA358632 Em:AI214048
Em:AI929470 Em:AA319216 Em:AA317625 Em:D81392 Em:AA317418
Em:AA366528 Em:C15186 Em:AI291270 Em:AA160660 Em:AA612605
Em:AW160832 Em:AA384405 Em:AA160657 Em:AA081744
Em:AI291126 Em:C15591 Em:AA384268 Em:AI929703 Em:AI302026
Em:W19547 Em:AI141727 Em:AA946933 Em:AI141583 Em:H61106
Em:W60581 Em:AI826617 Em:AA706545 Em:AA464771 Em:AA862327
Em:W68380 Em:AA641652 Em:AA384269 Em:AA527696 Em:W48778
Em:AA633989 Em:AI659486 Em:AW167301 Em:AI929428
Em:AW379411 Em:C15948 Em:AI745669 Em:AW024587 Em:AW379207;
match: proteins: Tr:Q9Y516 Tr:Q9Y517"
/evidence="not_experimental"
/product="dj79p11.1 (novel protein similar to mouse Bex2

```

```

(gene="brain-expressed X-linked protein 2))"
complement(30502. .32134)
/gene="dj79p11.1"
30507. .30988
/notes="match: STS: Em:G24641; match: STS: Em:T86927"
30508. .30852
/notes="match: STS: Em:R43117; match: STS: Em:G23964"
complement(30518. .30523)
/gene="dj79p11.1"
30541. .30922
/notes="match: STS: Em:H61107"
31543. .32262
/notes="CpG island"
/evidence="not_experimental"
31647. .31730
/notes="42 copies 2 mer cc 64% conserved"
complement(32655. .33316)
/notes="match: GSS: Em:AQ489303"
33027. .34514
/notes="LIME2 repeat: matches 4631. .6155 of consensus"
34528. .34786
/notes="LIME3A repeat: matches 4404. .4664 of consensus"
34787. .35085
/notes="AluSg repeat: matches 1. .300 of consensus"
35086. .35412

alignment_scores:
Quality: 468.00 Length: 129
Ratio: 4.000 Gaps: 2
Percent Similarity: 90.698 Percent Identity: 68.217

alignment_block:
US-09-327-750D-33 x AL133348/rev ..

Align seg 1/1 to reverse of: AL133348 from: 1 to: 40584

1 MetGluSerLysAspGlnGlyAla...LysAsnLeuAsnMetGluAsnAs 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
31132 ATGGAGTCCAAAGAGAACGCGCTTAACACATCTCATCTGTGGAAATGT 31083

16 pHISGlnLysLysGluGluLysGluGluLysProGlnAspThrIleLysA 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
31082 CAACACAGGAAATGATCAAAAGATGAAAGAGCAAGTTGCTAAATAAG 31033

33 rgGluProValAlaAlaProThrPheGluAlaGlyLysAsnCysAlaPro 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
31032 GGGAGCCCTTGGCCCTTACCT...TTGAATGTAGTCAATCTGTGCCT 30986

50 ArgGlyClyArgArgPheArgValArgGlnProIleSerHisTyrrAr 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
30985 AGAGGAATCGTAGGCGGTTCGCGTTAGCAGCCCATCTCTGCAGTATAG 30936

66 gTPAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluG 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
30935 ATGGGACATAATGCATAGGCTTGGAGAGCCACAGGCAAGGATCAGAGG 30886

83 luAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeu 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
30885 AGAATATGAAAGGATTTGGGAGGAGGTGAGACAGCTGATGGAAGCTG 30836

100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
30835 AGGCAAAACAGTTGAGTCATAGTTTGGGGGAGTCAGCAGTATCCCC 30786

116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
30785 TCACCATGACCATCAGCATGAGTTTGCCTTATGCCC 30749

seq_name: gb_htg:HSU8081

```

```

seq_documentation_block:
LOCUS HSU8081 41029 bp DNA HTG 10-JUL-2001
DEFINITION Homo sapiens chromosome X clone LL0XNC01-8081 map q22.1-22.3, ***

```



```

;
; TOPOLOGY: both
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 326..6277
US-08-836-325-9

alignment_scores:
  Quality: 72.50 Length: 58
  Ratio: 1.812 Gaps: 4
Percent Similarity: 68.966 Percent Identity: 36.207

alignment_block:
US-09-327-750D-34 x US-08-836-325-9/rev ..
Align seg 1/1 to reverse of: US-08-836-325-9 from: 1 to: 6452

25 LysAlaSerLysGlnSerGluGluGlu.....SerHisHisLeuG1 38
||||:||||:||||:||||:||||:||||:||||:||||:||||:
3220 AAGAGCCAGAAACAGAGTTCTAGGACCAAGAGTTCCCAATCACCATGACCA 3171

38 uGluValGluAsnLysLysProGlyGlyAsnValArgArgLysValArgA 55
:::||||:||||:||||:||||:||||:||||:||||:||||:
3170 TCATGTAACAAATGAAGCACATGCT.....CTGGCCCGGACCTC...C 3130

55 rgLeuValProAsnPhLeuTrpAlaIleProAsnArgHisValAspHis 71
:::||||:||||:||||:||||:||||:||||:||||:||||:
3129 ATGCAGTCCACATGCTCTCTATCCACTCTCCACACAGCAGCACTCGGAAC 3080

72 SerGluGlyGly...GluGluVal 78
:::||||:||||:||||:||||:||||:||||:||||:
3079 GATGAGGAGGAGTGGAAGAAGTC 3056

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-892-770-1

seq_documentation_block:
; Sequence 1, Application US/08892770
; Patent No. 5962670
; GENERAL INFORMATION:
; APPLICANT: Walling, Linda L.
; APPLICANT: Pautot, Veronique
; APPLICANT: Gu, Yong-Qiang
; APPLICANT: Chao, Wu Shaw
; TITLE OF INVENTION: Improved promoters for Enhancing Plant
; TITLE OF INVENTION: Productivity
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/892,770
; FILING DATE: 15-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Quine, Jonathan A.
; REGISTRATION NUMBER: P-41,261
; REFERENCE/DOCKET NUMBER: 023070-072100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 935 base pairs
;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..935
; OTHER INFORMATION: /note= "tomato acidic leucine
; OTHER INFORMATION: aminopeptidase 1 (LapAl) promoter
; OTHER INFORMATION: region"
US-08-892-770-1

alignment_scores:
  Quality: 72.00 Length: 60
  Ratio: 2.250 Gaps: 2
Percent Similarity: 53.333 Percent Identity: 28.333

alignment_block:
US-09-327-750D-34 x US-08-892-770-1 ..
Align seg 1/1 to: US-08-892-770-1 from: 1 to: 935

62 TrpAlaIleProAsnArgHisValAspHisSerGluGlyGlyGluGluVa 78
||||:||||:||||:||||:||||:||||:||||:||||:||||:
438 TGGTGTGTGTTCAATGAGATTATCGATCATATAAGAGAGGGGGAAGAT 487

78 lGlyArgPheValGlyGlnValMet.GluAlaLysArgHis..... 91
:::||||:||||:||||:||||:||||:||||:||||:||||:
488 GAACAGTTTTTTAATAAATATTTTGGCCCGCAGAAAAAACATTCACGCG 537

92 .....Serly 93
||||:||||:||||:||||:||||:||||:||||:||||:
538 GACTTTAAGTCATTACACTACCTGCCATTGCCATGACAGAGAAAATCGTCAA 587

93 sGluGlnGlnMetArgProTyrThrArg 102
||||:||||:||||:||||:||||:||||:||||:||||:
588 AAATGACACATGAGACCCCTACATAAGG 615

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-726-214-5

seq_documentation_block:
; Sequence 5, Application US/08726214
; Patent No. 610706
; GENERAL INFORMATION:
; APPLICANT: Tang, Wei-Jen
; APPLICANT: Gilman, Alfred G.
; TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,214
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/005,498
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSD:450

```

```

Percent Similarity: 59.000 Percent Identity: 20.000

alignment_block:
US-09-327-750D-34 x 5180810-3

Align seg 1/1 to: 5180810-3 from: 1 to: 1131

3 SerLysValLysGlnValIleLeuAspLeuThrValGluLysAspLysLys 19
:::||||:::
178 GCTAAACTTGTTCAGGTTGTTGAAACCAACACAGTTTGGAAAACGAA..... 222
19 sasnLysLysGlyGlyGlyLysGlnSerGluGluGluSerHisH 36
|||:::|||||:::
223 .....AAACTCAAGAGTGAAGATTGAGGAGAATAAGAAAA 256
36 isLeuGluValGluAsnLysLysProGlyGlyAsnValArgArgLys 52
:::||||:::
257 ATTAGCAAACTTAGCAAGAAATCAAGCA.....AAG 291
53 ValArgGluValProAsnPheLeuTrpAlaIleProAsnArgHisVa 69
:::||||:::
292 CTCGAAAAATTGGAGCTTGACTATCTCAAAAAATTAGATCAGCAGCAAA 341
69 laspHisSerGluGlyGlyGluGluValGlyArgPheValGlyGlnVal 86
:::||||:::
342 AGAGCACCAAAAGAACCAACAGAA.....C 367
86 etGluAlaLysArgHisSerLysGluGlnGlnMetArgProTyThrArg 102
:::||||:::
368 AAGAAGAGCGACAAAAAATCAAGAACAAATTAGAACGTAAATACCAACGA 417

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-045-583-6

seq_documentation_block:
; Sequence 6, Application US/09045583
; Patent No. 6287805
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6287805el Molecules of the G Prote
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045.583
; FILING DATE: 20-MAR-98
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36.207
; REFERENCE/DOCKET NUMBER: MNI-044
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1137 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1134
; US-09-045-583-6

alignment_scores:
  Quality: 71.00 Length: 114
  Ratio: 1.479 Gaps: 4
  Percent Similarity: 42.105 Percent Identity: 22.807

alignment_block:
  US-09-327-750D-34 x US-09-045-583-6/rev ..
  Align seg 1/1 to reverse of: US-09-045-583-6 from: 1 to: 1137
  4 LysVallYsGlnValIleLeuAspLeuThrValGluLysAsp..... 17
  876 CGTATAGTCCAGAGCGGTGGTGATCTCACAGTTCCTCCAAAGACATGCAGGT 827
  18 ....LysLysAsnLysLysGlyGlyLysAla.....SerLysGlnSerG 31
  826 CCAGCAACGAGGTGCAGAAACAAGGTGAGGTAATATGGGAACACACAGCATG 777
  31 luGluGluSerHis..... 35
  776 AAGAAACTATCACACAGGCGCGGCATCTCTCAGAGCCGCGCCCTGGCC 727
  35 ..... 35
  726 TGGCGGCTCAGCTGACCCAGAACGCAACCGATCGGGAGTAAAGAAGA 677
  36 .HisLeuGluGluValGluAsnLysLys..... 44
  676 TCATGGCCAGAGTGGGAGGAGAAACCCACAGAGTTTCAGTGGAGGCG 627
  45 .....ProGlycylAsnValArgArgLysValArgLysVal 57
  626 AGGTACAGCTTCCAATAGTGGTGGATGTCGCGCAAAATCCGATAGCAGTG 577
  58 ProAsnPheLeuTrpAlaIleProAsnArgHisValAspHis 71
  576 CCACACACCATCTAAGGTCTGGTGGATCTGCACAAAGACCAT 535
  seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-045-583-4

seq_documentation_block:
; Sequence 4, Application US/09045583
; Patent No. 6287805
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6287805el Molecules of the G Protein-Coupled
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,583
; FILING DATE: 20-MAR-98
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1134
; US-09-045-583-6

alignment_scores:
  Quality: 71.00 Length: 114
  Ratio: 1.479 Gaps: 4
  Percent Similarity: 42.105 Percent Identity: 22.807

alignment_block:
  US-09-327-750D-34 x US-09-045-583-6/rev ..
  Align seg 1/1 to reverse of: US-09-045-583-6 from: 1 to: 1137
  4 LysVallYsGlnValIleLeuAspLeuThrValGluLysAsp..... 17
  876 CGTATAGTCCAGAGCGGTGGTGATCTCACAGTTCCTCCAAAGACATGCAGGT 827
  18 ....LysLysAsnLysLysGlyGlyLysAla.....SerLysGlnSerG 31
  826 CCAGCAACGAGGTGCAGAAACAAGGTGAGGTAATATGGGAACACACAGCATG 777
  31 luGluGluSerHis..... 35
  776 AAGAAACTATCACCAGGCGCGGCATCTCAGAGCCGCGCCCTGGCC 727
  35 ..... 35
  726 TGGCGGCTCAGCTGACCCAGAACGCAACCGATCGGGAGTAAAGAAGA 677
  36 .HisLeuGluGluValGluAsnLysLys..... 44
  676 TCATGGCCAGAGTGGGAGGAGAAACCCACAGAGTTTCAGTGGAGGCG 627
  45 .....ProGlycylAsnValArgArgLysValArgLysVal 57
  626 AGGTACAGCTTCCAAATGGTGGCATGTCGCGCAAAATCCGATAGCAGTG 577
  58 ProAsnPheLeuTrpAlaIleProAsnArgHisValAspHis 71
  576 CCACACACCATCTAAGGTCTGGTGGATCTGCACAAAGACCAT 535
  seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-045-583-4

seq_documentation_block:
; Sequence 4, Application US/09045583
; Patent No. 6287805
; GENERAL INFORMATION:
; APPLICANT: Graham, Gerard J. et al.
; TITLE OF INVENTION: No. 6287805el Molecules of the G Protein-Coupled
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,583
; FILING DATE: 20-MAR-98
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

```

alignment_scores:
Quality: 71.00 Length: 100
Ratio: 1.203 Gaps: 3
Percent Similarity: 59.000 Percent Identity: 20.000

alignment_block:

US-09-327-750D-34 x 5180810-2 ..

Align seg 1/1 to: 5180810-2 from: 1 to: 1784

```
3 SerLysValLysGlnValLleLeuAspLeuThrValGluLysAspLysLy 19
505 GCTAACTTGTGAGTGTGTGAAACACACAAAGTTTGGAAACGAA..... 549
19 sAsnLysLysGlyGlyLysAlaSerLysGlnSerGluGluGluSerHisH 36
550 .....AAACTCAAGAGTGAGATGAGGAGCAATAAGAA 583
36 IsLeuGluGluValGluAsnLysLysProGlyGlyAsnValArgArgLys 52
584 ATTTAGACAAACTTAGCAAAAGAAATCAAGGA.....AAG 618
53 ValArgArgLeuValProAsnPheLeuTrpAlaLleProAsnArgHisVa 69
619 CTCGAAAAATGAGCTGTGACTATCTCAAAAAAATTAGATCAGCAGCACAA 668
69 LasPHisSerGluGlyGlyGluGluValGlyArgPheValGlyGlnValM 86
569 AGAGCACCAAAAGAACACAAAGAA.....C 694
86 etGluAlaLysArgHisSerLysGluGlnGlnMetArgProTyrThrArg 102
695 AAGAAGAGCGACAAAAAATCAAGAACAAATTAGACGCTAATAACCAACGA 744
```

seq_name: /cgn2.6/ptodata/2/lna/5B_COMB.seq:US-08-841-483-3

seq_documentation_block:

; Sequence 3, Application US/08841483B

; Patent No. 5976875

; GENERAL INFORMATION:

; APPLICANT: Prescott, Steven M.

; APPLICANT: Bunting, Michaeline

; APPLICANT: Tang, Wen

; APPLICANT: Topham, Matthew

; TITLE OF INVENTION: Diacylglycerol Kinase Isoforms Epsilon and Zeta and

; FILE REFERENCE: 2037.2.1a

; CURRENT APPLICATION NUMBER: US/08/841,483B

; EARLIER FILING DATE: 1997-04-22

; EARLIER FILING DATE: 1996-04-22

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; TYPE: DNA

; LENGTH: 3490

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (89)..(2875)

US-08-841-483-3

alignment_scores:
Quality: 71.00 Length: 82
Ratio: 1.651 Gaps: 4
Percent Similarity: 52.439 Percent Identity: 30.488

alignment_block:

US-09-327-750D-34 x US-08-841-483-3 ..

Align seg 1/1 to: US-08-841-483-3 from: 1 to: 3490

```
13 ThrValGluLysAspLysLysAsnLysLysGlyGly.....LysAl 26
|||||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
848 ACTCTGAAAGCAAGCAAGAAAGAAAGAGGGCATCTTCAAGAGGAAGTC 857
26 aSerLysGlnSerGluGluGluSerHis..... 35
|||||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
898 CAGCAAGAAAGGGCCTGAGGAGGGCGCTGGAGACCCCTTCATCATCAGGC 947
36 .....HisLeuGluGluValGluAsnLysLys 44
948 CCACCCCTCCCGCTCATGAAGCCCTGCTGGTGTGTTGTGAACCCCAAG 997
45 ProGlyGlyAsnValArgArgLysValArgArgLeuValProAsnPheLe 61
|||||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
998 AGTGGGGGCAACCAAGGTCGCAAGATC.....ATCCAGTCTTTCCT 1038
61 uTrpAlaLleProAsnArgHisVal...AspHisSerGluGlyGly 75
|||||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
1039 CTGGTATCTCAATCCCGCACAGTCTTCGACCTGAGCCAGGAGGG 1084
```

seq_name: /cgn2.6/ptodata/2/lna/6B_COMB.seq:US-09-382-911-3

seq_documentation_block:

; Sequence 3, Application US/09382911

; Patent No. 6221658

; GENERAL INFORMATION:

; APPLICANT: Prescott, Steven M.

; APPLICANT: Bunting, Michaeline

; APPLICANT: Tang, Wen

; APPLICANT: Topham, Matthew

; TITLE OF INVENTION: Diacylglycerol Kinase Isoforms Epsilon and Zeta and

; FILE REFERENCE: 2037.2.1a

; CURRENT APPLICATION NUMBER: US/09/382,911

; PRIOR FILING DATE: 1999-08-25

; PRIOR FILING DATE: 1997-04-22

; PRIOR FILING DATE: 1996-04-22

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; TYPE: DNA

; LENGTH: 3490

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (89)..(2875)

US-09-382-911-3

alignment_scores:

Quality: 71.00 Length: 82

Ratio: 1.651 Gaps: 4

Percent Similarity: 52.439 Percent Identity: 30.488

alignment_block:

US-09-327-750D-34 x US-09-382-911-3 ..

Align seg 1/1 to: US-09-382-911-3 from: 1 to: 3490

```
13 ThrValGluLysAspLysLysAsnLysLysGlyGly.....LysAl 26
|||||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
848 ACTCTGAAAGCAAGCAAGAAAGAAAGAGGGCATCTTCAAGAGGAAGTC 897
26 aSerLysGlnSerGluGluGluSerHis..... 35
|||||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
898 CAGCAAGAAAGGGCCTGAGGAGGGCGCTGGAGACCCCTTCATCATCAGGC 947
36 .....HisLeuGluGluValGluAsnLysLys 44
948 CCACCCCTCCCGCTCATGAAGCCCTGCTGGTGTGTTGTGAACCCCAAG 997
```



```

; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/807,043B
; FILING DATE: 199111212
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/764,364
; FILING DATE: 23-SEPTEMBER-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/728,838
; FILING DATE: 9-JULY-1991
; APPLICATION DATA:
; APPLICATION NUMBER: 07/705,702
; FILING DATE: 23-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5342774man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 253.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4698 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: singular
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; US-07-807-043B-5

alignment_scores:
    Quality: 71.00      Length: 93
    Ratio: 1.365        Gaps: 4
    Percent Similarity: 55.914      Percent Identity: 27.957

alignment_block:
US-09-327-750D-34 x US-07-807-043B-5/rev ..

Align seg 1/1 to reverse of: US-07-807-043B-5 from: 1 to: 4698

16 LysAspLysLysAsnLysGlyGlyLysAlaSer.LysGlnSerGluG 32
2134 AGGCAGGAAAGCAGCAGAAAGCGGGAGCCAGCGAGGGGAGGAC 2085
32 luGluSerHisLysLeuGluGluValGluAsnLys.....Lys 44
2084 AGGAAAGGCACACAGAGGGGAGCGGAGGGGGGATCTTGGAAA 2035
45 ProGlyGlyAsnValArgArgLysValArgArgLeuValProAsnPheLe 61
2034 CAGGTGCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1990
61 uTPAlaIleProAsnArgHisValAspHisSerGluGlyGluGluV 78
1989 .....CCAAACAACAAAAAGGTGAGCTGGAAGGTGGAGGGGAGC 1950
78 alGly.....ArgPheValGlyGlnValMetGluAlaLys 89
1949 CAGGAGGAAAGGCACAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGT 1900
90 ArgHisSerLysGluGlnGlnMetArg 98
1899 AGCGCTGCTAAAGTCGACGCAAGAGG 1873

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-299-849B-5

seq_documentation_block:
; Sequence 5, Application US/08299849B
; Patent No. 5612201
; GENERAL INFORMATION:
; APPLICANT: De Plaen, Etienne; Boon-Falleur, Thierry;
; APPLICANT: Leth , Bernard; Szikora, Jean-Pierre; De Smet, Charles;
; APPLICANT: Chomez, Patrick
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful In
; TITLE OF INVENTION: Determining Expression Of A Tumor Antigen Precursor
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/299,849B
; FILING DATE: 1-SEPTEMBER-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/037,230
; FILING DATE: 26-MARCH-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/04354
; FILING DATE: 22-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/807,043
; FILING DATE: 12-DECEMBER-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/764,364
; FILING DATE: 23-SEPTEMBER-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/728,838
; FILING DATE: 9-JULY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705,702
; FILING DATE: 23-May-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5612201man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5355
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4698 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; US-08-299-849B-5

alignment_scores:
    Quality: 71.00      Length: 93
    Ratio: 1.365        Gaps: 4
    Percent Similarity: 55.914      Percent Identity: 27.957

alignment_block:
US-09-327-750D-34 x US-08-299-849B-5/rev ..

Align seg 1/1 to reverse of: US-08-299-849B-5 from: 1 to: 4698

16 LysAspLysLysAsnLysGlyGlyLysAlaSer.LysGlnSerGluG 32
2134 AGGCAGGAAAGCAGCAGAAAGCGGGAGCGGAGGGGAGGAC 2085
32 luGluSerHisLysLeuGluGluValGluAsnLys.....Lys 44
2084 AGGAAAGGCACACAGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2035
45 ProGlyGlyAsnValArgArgLysValArgArgLeuValProAsnPheLe 61
2034 CAGGTGCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1990
61 uTPAlaIleProAsnArgHisValAspHisSerGluGlyGluGluV 78
1989 .....CCAAACAACAAAAAGGTGAGCTGGAAGGTGGAGGGGAGC 1950
78 alGly.....ArgPheValGlyGlnValMetGluAlaLys 89
1949 CAGGAGGAAAGGCACAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGT 1900
90 ArgHisSerLysGluGlnGlnMetArg 98
1899 AGCGCTGCTAAAGTCGACGCAAGAGG 1873

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-299-849B-5

seq_documentation_block:
; Sequence 5, Application US/08299849B
; Patent No. 5612201
; GENERAL INFORMATION:
; APPLICANT: De Plaen, Etienne; Boon-Falleur, Thierry;
; APPLICANT: Leth , Bernard; Szikora, Jean-Pierre; De Smet, Charles;
; APPLICANT: Chomez, Patrick
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful In
; TITLE OF INVENTION: Determining Expression Of A Tumor Antigen Precursor
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/299,849B
; FILING DATE: 1-SEPTEMBER-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/037,230
; FILING DATE: 26-MARCH-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/04354
; FILING DATE: 22-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/807,043
; FILING DATE: 12-DECEMBER-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/764,364
; FILING DATE: 23-SEPTEMBER-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/728,838
; FILING DATE: 9-JULY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705,702
; FILING DATE: 23-May-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5612201man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5355
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4698 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; US-08-299-849B-5
```

THIS PAGE BLANK (USPTO)

```
45 ProGlyGlyValArgArgLysValArgArgLeuValProAsnPheLe 61
   |||:||||: |||:||||:||||:||||: |||: |||: |||:
2034 CAAGGTGCRAAAAAAAGCAAAAGCAACCAACCAACCAACAA 1990
61 uTrpAlaIleProAsnArgHisValAspHisSerGlyGlyGluGluV 78
   |||:||||: |||:||||: |||: |||: |||: |||:
1989 .....CCAAACAAACAAAGGTGAGCTGGAAGGTGGAGGGGAGC 1950
78 aGly.....ArgPheValGlyGlnValMetGluAlaLys 89
   |||: |||: |||: |||: |||: |||: |||:
1949 CAGGAGGAAGGCACAAAAAGGGGAACGGGGCAGAAAGCAGGGAGAGGT 1900
90 ArgHisSerLysGluGlnGlnMetArg 98
   |||: |||: |||: |||: |||: |||: |||:
1899 AAGGCTGCTAAAGTCGAGCAAGAGG 1873
```


OM of: US-09-327-750D-34 to: N_Geneseq_1101:* out_format : pfs
Date: Mar 11, 2002 3:42 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09327750/runat_11032002_101154_20362/app_query.fasta_1.1472
-DB=N_Geneseq_1101 -OPMT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINWATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
-TRANS=human4.0.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=pfs
-NORM_ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09327750@cgn1_1.330 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-34

Query length: 118

Database: N_Geneseq_1101:*

Database sequences: 930621

Database length: 428662619

Search time (sec): 355.560000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517	+	372.00	800.02	865	1.7e-36
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13750	+	372.00	796.41	1229	2.8e-36
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03548	+	372.00	795.34	1364	3.2e-36
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF23529	+	171.50	358.04	891	7.2e-12
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAF21748	+	144.00	354.44	917	1.1e-11
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF23528	+	170.00	299.94	700	1.2e-08
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01004	+	143.00	301.51	485	1.0e-08
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01005	+	136.00	285.14	532	8.3e-08
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH75810	+	111.00	226.31	767	0.0002
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH58581	+	111.00	225.10	862	0.0002
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF59611	+	111.00	224.68	898	0.0002
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143	+	109.50	222.67	792	0.0003
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH03880	+	103.00	210.20	662	0.0012
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH03879	+	103.00	209.65	698	0.0013
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH160367	+	99.00	198.72	858	0.0054
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH10889	+	98.50	199.83	692	0.0047
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01003	+	92.50	203.23	401	0.0030
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAH23519	+	92.50	143.81	44453	6.20
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH20201	-	89.50	189.45	276	0.0178
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH15401	-	89.50	189.45	276	0.0178
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH15905	-	89.50	189.45	276	0.0178
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13308	+	87.50	180.85	415	0.0536
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH134661	+	87.50	180.85	415	0.0536
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03209	+	87.50	180.85	415	0.0536
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH7839	+	83.50	158.39	1566	0.9551
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH10225	+	83.50	114.98	250.27	0.9501
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAH08625	+	81.00	150.02	2069	2.80
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAH15484	+	80.50	159.18	762	0.8631
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH28355	-	80.00	114.62	52216	26.02
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH46153	+	78.00	137.19	3786	14.49
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH77225	+	78.00	136.98	3866	14.89
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH08642	+	78.00	136.82	3925	15.19
/SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:AAH086237	+	78.00	135.42	4500	18.19
/SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:AAH086236	+	77.50	135.31	4548	18.45
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH70222	+	77.50	131.37	5994	30.58
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAH39709	+	77.00	150.83	811	2.52
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH22174	+	77.00	147.68	1102	3.77
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH12793	+	76.00	156.05	394	1.29
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH46784	+	76.00	156.05	394	1.29
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH07190	+	76.00	156.05	394	1.29

/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH54129 + 76.00 133.47 23.35
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH93013 - 75.50 149.04 3.17
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH73887 - 75.50 142.27 7.55
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH72409 + 75.00 142.42 7.41
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH00010 - 75.00 133.35 23.71
289:
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517
seq_documentation_block:
ID AAH03517 standard; cDNA; 865 BP.
XX
AC AAH03517;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:352.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
full-length cDNAs defined in the specification, and for the detection
and/or diagnosis of the abnormality of the proteins encoded by the
full-length cDNAs.
XX
XX Claim 1; SEQ ID 352; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
full-length cDNAs defined in the specification. Where a primer set
comprises: (a) an oligo-dr primer and an oligonucleotide complementary
to the complementary strand of a polynucleotide which comprises one of
the 5602 nucleotide sequences defined in the specification, where the
oligonucleotide comprises at least 15 nucleotides; or (b) a combination
of an oligonucleotide comprising a sequence complementary to the
complementary strand of a polynucleotide which comprises a 5'-end
sequence and an oligonucleotide comprising a sequence complementary to a
polynucleotide which comprises a 3'-end sequence, where the
oligonucleotide comprises at least 15 nucleotides and the combination of
the 5'-end sequence/3'-end sequence is selected from those defined in
the specification. The primer sets can be used in antisense therapy and
in gene therapy. The primers are useful for synthesizing polynucleotides,
particularly full-length cDNAs. The primers are also useful for the
detection and/or diagnosis of the abnormality of the proteins encoded by
the full-length cDNAs. The primers allow obtaining of the full-length
cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
represent oligonucleotides, all of which are used in the exemplification
of the present invention.
XX
XX Sequence 865 BP; 239 A; 177 C; 224 G; 221 T; 4 other;
SQ

```
alignment_scores:
  Quality: 372.00      Length: 120
  Ratio: 3.758        Gaps: 2
  Percent Similarity: 82.500    Percent Identity: 60.000

alignment_block:
  US-09-327-750D-34 x AAH03517  ..

  Align seg 1/1 to: AAH03517 from: 1 to: 865

1  MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
   ||| ||||| : : : : : ||||| : : : : : ||||| : : : : :
191  ATGGAGTCCAAAGAGGAAGTACGCGGCAACAATCTCAACGGGGAATATGC 240

17  pLysLysAsnLysLysGlyGly...LysAlaSerLysGlnSerGluGluG 33
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
241  CCACAGAAACCAAGAGGGGAGCGAGCCGCCAGCAAGTAAAGAAAG 290

33  LuSerHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnVal 49
   ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
291  AATCCCGCATTTGGGAGGGGTCAAGGCCAGAAAGCCTGGAGGAATATC 340

50  ArgArg...LysValArgArgLeuValProAsnPheLeuTrpAlaIlePr 65
   ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
341  AGCGGGGGCGGAGTACGCGGATTTGCCCTAAATTTTCGATGGGCATACC 390

65  oAsnArgHisValAspHisSerGlyGlyGluGluValGlyArgPheV 82
   ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
391  TAATAGCATATTGAGCACATGAAAGCGAGAGATGATGTAGAAAGTTTG 440

82  alGlyGlnValMetGluAlaLysArgHisSerLysGluGlnMetArg 98
   ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
441  TAGGCAGATGATGGAATCAAGAGAAAGACTAGGGAACAGCAGATCAGG 490

99  ProTyrThrArgPheArgThrProGluProAsnHisTyrAspPheCy 115
   ||| ||||| : : : : : ||||| : : : : : ||||| : : : : :
491  CACTATATGCGCTTCCAACTCTCTGAACCTGACACCACTTATGACTTTTG 540

115  sLeuIlePro 118
   ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
541  CCTCATACCT 550
```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH13750

seq_documentation_block:

ID AAH13750 standard; cDNA; 1229 BP.

XX AC AAH13750;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA sequence SEQ ID NO:10656.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

XX DR Primer sets for synthesizing polynucleotides, particularly the 5602

XX PT full-length cDNAs defined in the specification, and for the detection

XX PT and/or diagnosis of the abnormality of the proteins encoded by the

XX PT full-length cDNAs

XX Claim 8; SEQ ID 10656; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to

CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

XX Sequence 1229 BP; 351 A; 243 C; 296 G; 339 T; 0 other;

alignment_scores:

Quality: 372.00 Length: 120

Ratio: 3.758 Gaps: 2

Percent Similarity: 82.500 Percent Identity: 60.000

alignment_block:

US-09-327-750D-34 x AAH13750 ..

Align seg 1/1 to: AAH13750 from: 1 to: 1229

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17

191 ATGGAGTCCAAAGAGGAAGTACGCGGCAACAATCTCAACGGGGAATATGC 240

17 pLysLysAsnLysLysGlyGly...LysAlaSerLysGlnSerGluGluG 33

241 CCACAGAAACCAAGAGGGGAGCGAGCCGCCAGCAAGTAAAGAAAG 290

33 LuSerHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnVal 49

291 AATCCCGCATTTGGGAGGGGTGAAGGCCAGCAAGCCTGGAGGAATATC 340

50 ArgArg...LysValArgArgLeuValProAsnPheLeuTrpAlaIlePr 65

341 AGCGGGGGCGGAGTACGCGGACTTGTCCCTAAATTTTCGATGGGCATACC 390

65 oAsnArgHisValAspHisSerGlyGlyGluGluValGlyArgPheV 82

391 TAATAGCATATTGAGCACAAATCAAGCGAGAGATGATGTAGAAAGTTTG 440

82 alGlyGlnValMetGluAlaLysArgHisSerLysGluGlnMetArg 98

441 TAGGCAGATGATGGAATCAAGAGAAAGACTAGGGAACAGCAGATCAGG 490

99 ProTyrThrArgPheArgThrProGluProAsnHisTyrAspPheCy 115

491 CACTATATGCGCTTCCAACTCTCTGAACCTGACACCACTTATGACTTTTG 540

Quality:	372.00	Length:	120
Ratio:	3.758	Gaps:	2
Percent Similarity:	82.500	Percent Identity:	60.000


```
XX
SQ Sequence 662 BP; 177 A; 144 C; 192 G; 137 T; 12 other;

alignment_scores:
  Quality: 103.00      Length: 132
  Ratio: 1.411        Gaps: 6
  Percent Similarity: 55.303      Percent Identity: 30.303

alignment_block:
US-09-327-750D-34 x AAC03880 ..
Align seg 1/1 to: AAC03880 from: 1 to: 662
1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrVal..... 14
207 ATGGAGTCCAAAGAGGAAACGAGGCTTAAACAATCTCATCTGGAAATGT 256
15 .....GluLysAspLysLysAsnLysLysGlyGlyLysAlaSerLysG 29
257 CAACGAGAAATCATGAAAGATGAAAGAGGACCAAGTGWCTAATAAA.. 305
29 lnSerGluGluGluSerHisHisLeuGluGluValGluAsnLysLysPro 45
306 ..GGGAGCCCTTGGCCCTACCTTGTGCTGTTGTGTTGTGCTGCT 353
46 GlyGlyAsnValArgArg....LysValArgArgLeuValProAsnPheLe 61
354 AGAGGAATTCGTAGGCGGTTCCGCGTTAGGACGCCCATCTGCGAGTATAG 403
61 uTrpAlaIleProAsnArgHisValAsp.....HisSerG 73
404 ATGGGATATGATCATGATGGCTTGGAGACCAACGAGCAAGGATGANAGAAG 453
73 luGlyGlyGluGluValGlyArgPheValGlyGlnValMetGluAlaLys 89
454 AGAATATGGAAGGATTGGGAGGAGGTGACACAGCTGATGGAAAGCTG 503
90 ArgHisSerLysGluGlnGlnMetArgProTyrThrArgPheArgThrPr 106
504 AGG.....GAAAGACAGTTGAGTCATAGTCTGCGGGCAGTCAGCAC 544
106 oGluPro.....AspAsnHisTyrAspPheCysLeuIlePro 118
545 TGACCCCTCCATCACCATGACCATCATGATGATGATGATGATGATGATG 590
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAC03879
seq_documentation_block:
ID AAC03879 standard; cDNA; 698 BP.
XX
AC AAC03879;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 3877.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
```

```
XX
DR WPI: 2000-500381/45.
P-PSDB; AAG03873.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 3877; 7lpp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 698 BP; 187 A; 151 C; 210 G; 139 T; 11 other;

alignment_scores:
  Quality: 103.00      Length: 132
  Ratio: 1.411        Gaps: 6
  Percent Similarity: 55.303      Percent Identity: 30.303

alignment_block:
US-09-327-750D-34 x AAC03879 ..
Align seg 1/1 to: AAC03879 from: 1 to: 698
1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrVal..... 14
243 ATGGAGTCCAAAGAGGAAACGAGGCTTAAACAATCTCATCTGCGAAATGT 292
15 .....GluLysAspLysLysAsnLysLysGlyGlyLysAlaSerLysG 29
293 CAACGAGAAATGATGAAAGATGAAAGAGGCAAGTGWCTAATAAA.. 341
29 lnSerGluGluGluSerHisHisLeuGluGluValGluAsnLysLysPro 45
342 ..GGGAGCCCTTGGCCCTACCTTGTGATGTYTRGTGAATACTGTGTCCT 389
46 GlyGlyAsnValArgArg....LysValArgArgLeuValProAsnPheLe 61
390 AGAGGAATTCGTAGGCGGTTCCGCGTTAGGACGCCCATCTGCGAGTATAG 439
61 uTrpAlaIleProAsnArgHisValAsp.....HisSerG 73
440 ATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 489
73 luGlyGlyGluGluValGlyArgPheValGlyGlnValMetGluAlaLys 89
490 ACAATATGGAAGGATTGGGAGGAGGTGACACAGCTGATGCAAGAGCTG 539
90 ArgHisSerLysGluGlnGlnMetArgProTyrThrArgPheArgThrPr 106
540 AGG.....GAAAGCAGTTGAGTCATAGTCTGCGGGCAGTCAGCAC 580
106 oGluPro.....AspAsnHisTyrAspPheCysLeuIlePro 118
581 TGACCCCTCCATCACCATGACCATCATGATGATGATGATGATGATGATG 626
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAI60367
seq_documentation_block:
```

THIS PAGE BLANK (USPTO)

Sequence	Strd	Orig	zScore	EScore	Len	Documentation
gb_pr:AK000959	+	372.00	661.44	1.4e-28	1269	! AX000959 Homo sapiens cDNA FLJ
gb_pat:AX100231	+	372.00	660.55	1.6e-28	1364	! AX100231 Sequence 32 from Paten
gb_pr:HS035G19	+	372.00	627.19	1.5e-26	89648	! AL035494 Human DNA sequence fr
gb_pr:HM00C	+	171.50	302.23	1.2e-08	691	! M38188 Human unknown protein fr
gb_pr:AF187064	+	171.50	302.23	1.5e-08	891	! AF187064 Homo sapiens p75NTR-as
gb_pr:BC003190	+	170.00	300.51	1.8e-08	793	! BC003190 Homo sapiens, p75NTR-a
gb_to:AF187065	+	148.00	264.39	1.9e-06	519	! AF187065 Rattus norvegicus p75N
gb_sts:G72708	-	145.00	260.75	3.0e-06	421	! G72708 MRC 4953-4954:99139031
gb_to:AF187066	+	145.00	256.44	5.2e-06	700	! AF187066 Mus musculus p75NTR-as
gb_ro:AF097440	+	145.00	254.75	6.5e-06	854	! AF097440 Mus musculus brain exp
gb_pr:HSV351F8	+	138.00	208.36	0.0025	45678	! 270719 Human DNA sequence fr
gb_sts:G35294	-	118.50	211.86	0.0016	477	! G35294 human STS SHGC-37409, se
gb_to:AF097439	+	116.50	204.02	0.0044	785	! AF097439 Mus musculus brain exp
gb_pr:AF163416	+	112.50	198.50	0.0088	642	! AF183416 Homo sapiens ovarian ex
gb_to:AF201819	+	112.50	196.35	0.0117	828	! AF201819 Homo sapiens uncharac
gb_pr:AF1598P4	-	112.50	164.41	0.6998	35714	! AL008708 Human DNA sequence fr
gb_pr:AF251053	+	111.00	194.55	0.0147	744	! AF251053 Homo sapiens X-linked
gb_pat:AX078272	+	111.00	192.95	0.0180	898	! AX078272 Sequence 76 from Paten
gb_pr:HSV870H8	+	111.00	162.82	0.8586	31321	! 270233 Human DNA sequence fr
gb_pr:AL133348	-	111.00	160.62	1.14	40584	! AL133348 Human DNA sequence fr
gb_btg:HSU80B1	+	111.00	160.53	1.15	41029	! AL022169 Homo sapiens chromos
gb_pr:AF237783	+	109.50	191.32	0.1122	791	! AF237783 Homo sapiens brain-exp
gb_ro:AF097438	+	105.00	182.74	0.0268	835	! AF097438 Mus musculus brain-exp
gb_to:AF051347	+	105.00	182.71	0.0670	838	! AF051347 Mus musculus REX-3 mRN
gb_to:AF097437	+	105.00	174.26	0.1981	2269	! AF097437 Mus musculus Bex1 pr
gb_pr:HS714B7	+	102.50	137.78	21.32	98274	! 299755 Human DNA sequence fr
gb_ba:AF071813	+	96.00	163.24	0.8138	1225	! AF071813 Streptococcus pneumon
gb_btg:AC023561	+	95.50	120.50	195.44	169808	! AC023561 Homo sapiens chrom
gb_pr:AC074348	+	95.00	120.24	202.20	157511	! AC074348 Homo sapiens clone
gb_btg:AC048344	+	95.00	119.13	232.92	179380	! AC048344 Homo sapiens chrom
gb_btg:AC023341	+	94.50	118.72	245.66	169365	! AC023341 Homo sapiens chrom
gb_pr:HSAC000373	+	94.00	130.14	56.75	39602	! AC000373 Human Cosmid g1862a2
gb_pr:HSAC000367	+	94.00	129.38	62.62	43349	! AC000367 Human Cosmid g1862a2
gb_btg:AC012232	+	94.00	117.57	284.63	174337	! AC012232 Homo sapiens clone
gb_pr:HS753P9	+	92.50	118.28	259.95	116561	! AC023653 Human DNA sequence
gb_pr:AL135809	+	92.50	118.01	269.13	120340	! AL135809 Human DNA sequence
gb_pr:AF195953	+	92.50	113.34	489.70	208618	! AF195953 Homo sapiens membra
gb_in:CSA7431	+	92.00	151.74	3.56	2029	! A0007431 Cupienius salei mRNA
gb_btg:AL512842	+	92.00	114.96	397.64	154889	! AL512842 Homo sapiens chrom
gb_btg:AL091457	+	92.00	112.30	559.73	212080	! AC091457 Mus musculus clone

```

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
111 ||||| : : : : : ||||| : : : : :
191 ATGGAGTCCAAAGAGAACTAGCGGCAACAATCTCAACGGGGAATGC 240

17 pLysLysAsnLysLysGlyGly...LysAlaSerLysGlnSerGluGluG 33
: : : : : : : : : : : : : : : : : : : : : : : : : : :
241 CCAACAAGAAACGAAGAGGAGGAGCGAGCCGCCACCGAGAATGAAGAAG 290.

33 LuSerHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnVal 49
||||| : : : : : : : : : : : : : : : : : : : : :
291 ATCCCGCGCATTTGGAGGGGTGAGGCGCAAGCCTGGAGGAATATC 340

50 ArgArg...LysValArgArgLeuValProAsnPheLeuTrpAlaIlePr 65
||||| : : : : : : : : : : : : : : : : : : : : :
341 AGCGGGGGCGAGTTAGCGCACTTGCTCCCTAATTTTCGATGGCCATACC 390

65 oAsnArgHisValAspHisSerGluGlyGlyGluGluValGlyArgPheV 82
||||| : : : : : : : : : : : : : : : : : : : : :
391 TATAGGCATATTGAGCACAATCAAGCGAGAGATGATGTAGAAAGTTTG 440

82 aGlyGlnValMetGluAlaLysArgHisSerLysGluGlnGlnMetArg 98
||||| : : : : : : : : : : : : : : : : : : : : :
441 TAGGCAGATGATGGAATCAAGAGAAAGACTAGGGAACAGCAGATGAGG 490

99 ProTyrThrArgPheArgThrProGluProAspAsnHisTyrAspPheCy 115.
||| ||||| : : : : : : : : : : : : : : : : : : : :
491 CACTATATGCGCTTCCAAACTCTGAACTGACACCATATGACTTTTG 540

115 sleullePro 118
|||||
541 CCTCATACCT 550

```

seq_name: gb_pat:AX100231

seq_documentation_block:

LOCUS AX100231 1364 bp DNA PAT 02-APR-2001
 DEFINITION Sequence 32 from Patent WO0119860.
 ACCESSION AX100231
 VERSION AX100231.1 GI:13539115
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 1364)
 Tang,Y.T., Hillman,J.L., Yue,H., Reddy,R., Lal,P., Shah,P.,
 Azimzai,Y., Baughn,M.R., Lu,D.A., Bandman,O., Shih,L.L. and
 Patterson,C.
 TITLE Proteins associated with cell differentiation
 JOURNAL Patent: WO 0119860-A 32 22-MAR-2001;
 Incyte Genomics Inc. (US)
 FEATURES
 source
 1..1364
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 1990956CB1"

BASE COUNT 411 a 269 c 322 g 362 t
 ORIGIN

alignment_scores:

Quality: 372.00 Length: 120
 Ratio: 3.758 Gaps: 2
 Percent Similarity: 82.500 Percent Identity: 60.000

alignment_block:

US-09-327-750D-34 x AX100231 ..

Align seg 1/1 to: AX100231 from: 1 to: 1364

```

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
111 ||||| : : : : : ||||| : : : : :
250 ATGGAGTCCAAAGAGAACTAGCGGCAACAATCTCAACGGGGAATGC 299

```

```

17 pLysLysAsnLysLysGlyGly...LysAlaSerLysGlnSerGluGluG 33
: : : : : : : : : : : : : : : : : : : : : : : : : : :
300 CCAACAAGAAACGAAGAGGAGGAGCGAGCCGCCACCGAGAATGAAGAAG 349

33 LuSerHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnVal 49
||||| : : : : : : : : : : : : : : : : : : : : :
350 AATCCCGCGCATTTGGAGGGGTGAGGCGCAAGCCTGGAGGAATATC 399

50 ArgArg...LysValArgArgLeuValProAsnPheLeuTrpAlaIlePr 65
||||| : : : : : : : : : : : : : : : : : : : : :
400 AGCGGGGGCGAGTTAGCGCACTTGCTCCCTAATTTTCGATGGCCATACC 449

65 oAsnArgHisValAspHisSerGluGlyGlyGluGluValGlyArgPheV 82
||||| : : : : : : : : : : : : : : : : : : : : :
450 TATAGGCATATTGAGCACAATCAAGCGAGAGATGATGTAGAAAGTTTG 499

82 aGlyGlnValMetGluAlaLysArgHisSerLysGluGlnGlnMetArg 98
||||| : : : : : : : : : : : : : : : : : : : : :
500 TAGGCAGATGATGGAATCAAGAGAAAGACTAGGGAACAGCAGATGAGG 549

99 ProTyrThrArgPheArgThrProGluProAspAsnHisTyrAspPheCy 115
||| ||||| : : : : : : : : : : : : : : : : : : : :
550 CACTATATGCGCTTCCAAACTCTGAACTGACACCATATGACTTTTG 599

115 sleullePro 118
|||||
600 CCTCATACCT 609

```

seq_name: gb_pr:HS635G19

seq_documentation_block:

LOCUS HS635G19 69648 bp DNA PRI 23-NOV-1999
 DEFINITION Human DNA sequence from clone 635G19 on chromosome Xq22.1-22.3
 Contains a LAMR1 (Laminin Receptor 1 (67kD) (RPSA), 40S Ribosomal
 Protein SA, P40)) pseudogene and part of a novel protein. Contains
 ESTs and GSSs, complete sequence.

ACCESSION AL035494

VERSION AL035494.8 GI:4775622

KEYWORDS HTG: 40S Ribosomal Protein P40; 40S Ribosomal Protein SA; CpG
 Island; Laminin Receptor 1; LAMR1; RPSA.

SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 69648)

REFERENCE

AUTHORS Bird,C

TITLE Direct Submission

JOURNAL

Submitted (02-JUN-1999), Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk

COMMENT

On May 11, 1999 this sequence version replaced gi:4678749.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.

The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:

Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL

This sequence has been finished according to sequence map criteria
 as follows. An attempt is made to resolve all sequencing problems,
 such as compressions and repeats, but not necessarily within known
 annotated human repeat sequence elements (e.g. Alu). Where the
 sequence is ambiguous, there is an annotation using the 'unsure'
 feature key.

This sequence was generated from part of bacterial clone contigs of
 human chromosome X, constructed by the Sanger Centre Chromosome X
 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/ChrX
 635G19 is from the library RPC14 constructed at the Roswell Park
 Cancer Institute by the group of Pieter de Jong. For further

details see [http://bacpac.med.buffalo.edu/ VECTOR:pcypac2](http://bacpac.med.buffalo.edu/VECTOR:pcypac2)
 IMPORTANT: This sequence is not the entire insert of clone 635G19.
 It may be shorter because we only sequence overlapping sections
 once, or longer because we arrange for a small overlap between
 neighbouring submissions.
 The true left end of clone U65A4 (281014) is at 69549 in this
 sequence. The true right end of clone U101D3 (285997) is at 100 in
 this sequence.

FEATURES

Location/Qualifiers
 source

```

1..69648
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q22.1-22.3"
/clone="RP4-635G19"
/clone.lib="RPCI-4"
439..483
repeat_region
/feature="TAR1 repeat: matches 707..749 of consensus"
501..661
repeat_region
/feature="AluYb repeat: matches 137..311 of consensus"
662..897
repeat_region
/feature="118 copies 2 mer aa 70% conserved"
915..1331
repeat_region
/feature="MSTA repeat: matches 1..419 of consensus"
1334..1517
repeat_region
/feature="92 copies 2 mer ga 77% conserved"
1927..2020
repeat_region
/feature="MIR repeat: matches 109..217 of consensus"
2066..2262
repeat_region
/feature="L1ME2 repeat: matches 5825..6022 of consensus"
2263..2574
repeat_region
/feature="AluSp repeat: matches 1..313 of consensus"
2575..2819
repeat_region
/feature="L1ME2 repeat: matches 5583..5825 of consensus"
2821..3039
repeat_region
/feature="Alu repeat: matches 1..311 of consensus"
3104..3384
repeat_region
/feature="L1ME2 repeat: matches 5304..5591 of consensus"
3391..4007
repeat_region
/feature="L1MB1 repeat: matches 5558..6162 of consensus"
4012..4137
repeat_region
/feature="Trig2a repeat: matches 11..123 of consensus"
4141..4252
repeat_region
/feature="L1 repeat: matches 2920..3031 of consensus"
4265..4350
repeat_region
/feature="HY4 repeat: matches 1..87 of consensus"
4350..4678
repeat_region
/feature="MER74A repeat: matches 28..369 of consensus"
4914..5095
repeat_region
/feature="MER5A repeat: matches 8..189 of consensus"
5224..5555
repeat_region
/feature="L2 repeat: matches 2423..2746 of consensus"
6040..6086
repeat_region
/feature="L2 repeat: matches 1676..1720 of consensus"
7413..7478
repeat_region
/feature="Alu repeat: matches 239..304 of consensus"
7628..7928
repeat_region
/feature="AluYb8 repeat: matches 1..316 of consensus"
7952..7987
repeat_region
/feature="18 copies 2 mer aa 81% conserved"
7990..8305
repeat_region
/feature="AluJo repeat: matches 1..309 of consensus"
8436..8538
repeat_region
/feature="MIR repeat: matches 99..212 of consensus"
10836..11442
repeat_region
/feature="L1MD repeat: matches 1357..1972 of consensus"
11443..11571
repeat_region
/feature="FLAM.A repeat: matches 2..129 of consensus"
11572..11946
repeat_region
/feature="L1MD repeat: matches 977..1357 of consensus"
11947..12245
repeat_region
/feature="AluY repeat: matches 1..301 of consensus"
12246..13098
repeat_region
/feature="L1MD repeat: matches 275..977 of consensus"
13094..13223
repeat_region
/feature="L1MD repeat: matches 17..113 of consensus"
13523..13820
repeat_region
/feature="MLTID repeat: matches 177..505 of consensus"
13823..13932
repeat_region
/feature="55 copies 2 mer ct 77% conserved"
13936..13977
repeat_region
/feature="MLTID repeat: matches 201..241 of consensus"
13978..14240
repeat_region
/feature="AluSg1 repeat: matches 1..270 of consensus"
14245..14534
repeat_region
/feature="AluSc repeat: matches 1..290 of consensus"
14535..14721
repeat_region
/feature="MLTID repeat: matches 2..201 of consensus"
14724..15604
repeat_region
/feature="L1MC/D repeat: matches 4646..5565 of consensus"
15605..15909
repeat_region
/feature="AluX repeat: matches 24..312 of consensus"
16069..16363
repeat_region
/feature="AluX repeat: matches 1..291 of consensus"
16482..16641
repeat_region
/feature="AluYb repeat: matches 136..295 of consensus"
16671..16953
repeat_region
/feature="AluX repeat: matches 1..284 of consensus"
17164..17280
repeat_region
/feature="L1P repeat: matches 4499..4619 of consensus"
17326..17377
repeat_region
/feature="26 copies 2 mer at 79% conserved"
17704..17863
repeat_region
/feature="AluSc repeat: matches 118..295 of consensus"
17864..18175
repeat_region
/feature="AluYb8 repeat: matches 1..318 of consensus"
18176..18298
repeat_region
/feature="AluSc repeat: matches 1..118 of consensus"
18848..18873
repeat_region
/feature="13 copies 2 mer tg 100% conserved"
18876..19171
repeat_region
/feature="AluSg repeat: matches 1..295 of consensus"
19183..19232
repeat_region
/feature="25 copies 2 mer ac 100% conserved"
19246..19533
repeat_region
/feature="AluY repeat: matches 5..292 of consensus"
19539..19837
repeat_region
/feature="MER74B repeat: matches 61..368 of consensus"
20016..20194
repeat_region
/feature="MER74A repeat: matches 295..487 of consensus"
20394..20692
repeat_region
/feature="AluSg repeat: matches 1..302 of consensus"
20701..20809
repeat_region
/feature="LTR1 repeat: matches 682..785 of consensus"
20877..21061
repeat_region
/feature="LTR1 repeat: matches 1..195 of consensus"
21186..21636
repeat_region
/feature="LTR29 repeat: matches 3..498 of consensus"
21609..21737
repeat_region
/feature="MER34 repeat: matches 409..543 of consensus"
21771..22076
repeat_region
/feature="AluY repeat: matches 2..308 of consensus"
22878..22974
repeat_region
/feature="MIR repeat: matches 92..191 of consensus"
23452..23582
repeat_region
/feature="FLAM.C repeat: matches 1..131 of consensus"
23583..23624
repeat_region
/feature="21 copies 2 mer ta 76% conserved"
23735..24041
repeat_region
/feature="AluSg1 repeat: matches 1..308 of consensus"
24045..24072
repeat_region
/feature="14 copies 2 mer tt 89% conserved"
25274..25582
repeat_region
/feature="AluYb repeat: matches 1..289 of consensus"
25583..25747
misc_feature
/feature="match: GSS B63622"

```


617 CCATCATGATGAATTTGCTTATGCGCT 644

seq_name: gb_pr:AF187064

seq_documentation_block:

LOCUS AF187064 891 bp mRNA PRI 11-JUN-2000

DEFINITION Homo sapiens p75NTR-associated cell death executor (NADE) mRNA,

complete cds.

ACCESSION AF187064

VERSION AF187064.1 GI:8452893

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 891)

Mukai.J., Hachiya.T., Shoji-Hoshino.S., Kimura.M.T., Nadano.D.,

Suvanto.P., Hanaoka.T., Li.Y., Irie.S. and Sato.T.A.

NADE, a p75NTR-associated cell death executor, is involved in

signal transduction mediated by the common neurotrophin receptor

p75NTR

J. Biol. Chem. 275 (23), 17566-17570 (2000)

20298829

2 (bases 1 to 891)

Mukai.J., Hachiya.T., Hoshino.S., Kimura.M., Nadano.D., Suvanto.P.,

Hanaoka.T., Li.Y., Irie.S. and Sato.T.A.

Direct Submission

Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia

University, 630 West 168th St., P&S 11-451, New York, NY 10032, USA

Location/Qualifiers

1..891

/organism="Homo sapiens"

/db_xref="taxon:9606"

1..891

/gene="NADE"

312..647

/gene="NADE"

/note="involved in the common neurotrophin receptor

p75NTR-mediated signal transduction"

/codon_start=1

/product="p75NTR-associated cell death executor"

/protein_id="AAF5129.1"

/db_xref="GI:8452894"

/translation="MANIQNEEMEQPMQNGEDRPLGGGEGHQAGNRRGQARRLA

PNRWAI PNRIQINDGMGSDGDDMEIFEMREIRKRLRLRLNCLRLMGLSNNHHD

HHDFCLMP"

BASE COUNT 251 a 182 c 224 g 234 t

ORIGIN

alignment_scores:

Quality: 171.50 Length: 126

Ratio: 2.257 Gaps: 3

Percent Similarity: 60.317 Percent Identity: 32.540

alignment_block:

US-09-327-750D-34 x AF187064

Align seg 1/1 to: AF187064 from: 1 to: 891

12 LeuThrValGluLysAspLysLysAsnLysLysGlyGlyLysAlaSerIy 28

267 TTACTTGTTCAAGAAAACACCAAGAAAATAATCTCATCTGCG 316

28 sGlnSerGluGluSerHisLeuGluGluVal..... 40

317 AAATATTACCAAGAAAACCAAGAGATGGAGCAGCTATGCAGAAATCGAG 366

41GluAsnLysLysProGlyGlyAsn 48

367 AGGAAGACCCCTTTGGGAGGAGGTGAAGGCCACCGCTGCAGGAAT 416

49 ValArgArgLysValArgArgLeuValProAsnPheLeuTrpAlaIlePr 65

417 CGACGGGACAGGCTCGCGACTTGCCTTAATTTTCGATGGCCATACC 466

65 oAsnArgHisValAspHisSerGluGly.....GlyGluGluValGlyA 80

467 CAATAGCGCAATCAATGATGGGATGGTGGAGATGGATGATATGGA 516

80 rgPheValGlyGlnValMetGluAlaLysArgHisSerLysGluGln 96

517 TATTCATGGAGGAGAGAGAAATCAGAAGAACTTAGGGAGCTGCAG 566

97 MetArgProTyrThrArgPheArgThrProGluProAspAsnHisTyrAs 113

567 TTGAGGAATTTGTGCGTATCTTATGGGAGGCTCTCTAATCACCATGA 616

113 p.....PheCysLeuIlePro 118

617 CCATCATGATGAATTTTGCCTTATGCGCT 644

seq_name: gb_pr:BC003190

seq_documentation_block:

LOCUS BC003190 793 bp mRNA PRI 12-JUL-2001

DEFINITION Homo sapiens, p75NTR-associated cell death executor; ovarian

granulosa cell protein (13kD), clone MGC:802 IMAGE:3357965, mRNA,

complete cds.

ACCESSION BC003190

VERSION BC003190.1 GI:13112030

KEYWORDS MGC.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 793)

Strausberg,R.

Direct Submission

Submitted (13-FEB-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov

Contact: MGC help desk

Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcqsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,

Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,

Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo

Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven

Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline

Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,

Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,

George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 6 Row: k Column: 22

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7657043.

Location/Qualifiers

1..793

/organism="Homo sapiens"

/db_xref="LocusID:27018"

/db_xref="taxon:9606"

/clone="MGC:802 IMAGE:3357965"

/tissue_type="Eye, retinoblastoma"

/clone_lib="NTH_MGC_16"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

FEATURES

source

STS 08-AUG-2001
e spleen Sus scrofa STS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus. 1 (bases 1 to 421)
 Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L. and Keele, J.W.
 Single nucleotide polymorphism (SNP) discovery in expressed porcine genes
 unpublished (2001)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Contact: Freking BA
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4278
 Fax: 402 762 4173

Email: freking@email.marc.usda.gov
 Primer A: CCAATGGGTGAAACTCTACT
 Primer B: GCAGCAGCAATAGACG
 STS size: 500

PCR Profile:

Hotstart: 95 degrees for 15 minutes
 Denature: 95 degrees for 30 seconds
 Anneal: 56 degrees
 Extension: 68 degrees for 2 minutes
 Cycles: 32 to 45

Protocol:
 Template: 50-200 ng genomic DNA
 Primer: each 20 pmoles
 dNTPs: each 88 uM
 Taq Polymerase: 0.25 units (Qiagen HotStar)

Buffer: Commercially supplied Qiagen HotStar buffer

The STS is derived from PCR amplicons generated from genomic DNA, sequenced from each end using the amplification primers. The sequence does not necessarily represent the entire amplicon. Sequence derived from Polyphred was trimmed from each end of each unique contig until five consecutive bases exceeded a quality score threshold of 20, and the next 10 bases averaged a quality score of 20 or greater. Amplicon size was estimated by agarose gel electrophoresis.

FEATURES
 source

Location/Qualifiers
 1..421
 /organism="Sus scrofa"
 /strain="white composite, duroc, meishan, minzhu, fengjing, crossbreds"
 /db_xref="taxon:9823"
 /sex="male and female"
 /clone_lib="SCF - porcine spleen"
 /dev_stages="adult"
 /note="Organ: spleen"
 <1..>421

BASE COUNT 130 a 101 c 86 g 102 t 2 others
 ORIGIN

alignment_scores:
 Quality: 145.00 Length: 77
 Ratio: 2.843 Gaps: 2
 Percent Similarity: 66.234 Percent Identity: 40.260

alignment_block:

US-09-327-750D-34 x G72708/rev ..

Align seg 1/1 to reverse of: G72708 from: 1 to: 421

48 AsnValArgLysValArgArgLeuValProAsnPhelTrpAlaIrl 64

421 AATAGACGGGCAAGCTGCCCACTTGCCCTAATTCGATGGCCCAT 372

64 eProAsnArgHisValAspHisSerGluGly.....GlyGluValG 79

371 ACCCAATAGCAGATCAATGATGGGATGGTGGAGATGGATGATGG 322

79 lyArgPheValGlyGlnValMetGluAlaLysArgHisSerLysGluGln 95
 321 AAATGTTTCATGGAGGAGATGAGAGAAATCAGGAGAAACTTAGGGAGCTG 272
 96 GlnMetArgProTyrThrArgPheArgThrProGluProAspAsnHisTy 112
 271 CAGTTGAGGAATTGTCTGCGTATCTCTTATGGGGAGCTCTCTAATCACCA 222
 112 rAsp.....PheCysLeuIlePro 118
 221 TGACCATCATGATGAATTTTGCTTATGCCT 191

seq_name: gb_ro:AF187066

seq_documentation_block: 700 bp mRNA ROD 11-JUN-2000
 LOCUS AF187066
 DEFINITION Mus musculus p75NTR-associated cell death executor (Nade) mRNA,
 complete cds.

ACCESSION AF187066
 VERSION AF187066
 KEYWORDS AF187066.1 GI:8452897

SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 700)
 AUTHORS Mukai,J., Hachiya,T., Shoji-Hoshino,S., Kimura,M.T., Nadano,D., Suvanto,P., Hanaoka,T., Li,Y., Irie,S. and Sato,T.
 TITLE Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia University, 630 West 168th St., P&S 11-451, New York, NY 10032, USA

J. Biol. Chem. 275 (23), 17566-17570 (2000)
 MEDLINE 20298829
 REFERENCE 2 (bases 1 to 700)
 AUTHORS Mukai,J., Hachiya,T., Hoshino,S., Kimura,M., Nadano,D., Suvanto,P., Hanaoka,T., Li,Y., Irie,S. and Sato,T.
 TITLE Direct Submission
 JOURNAL Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia University, 630 West 168th St., P&S 11-451, New York, NY 10032, USA

FEATURES Location/Qualifiers
 1..700

/organism="Mus musculus"
 /strain="BALB/c"
 /db_xref="taxon:10090"

gene 1..700

CDS 177..551

/gene="Nade"
 /note="Involved in the common neurotrophin receptor p75NTR-mediated signal transduction; NADE"
 /codon_start=1
 /product="p75NTR-associated cell death executor"
 /protein_id="AAF75131.1"
 /db_xref="GI:8452898"
 /translation="MANVQNEEMEOPLQNGOEDRPVGGEGHQPAAANNHNNHNNHHRGQARLAPNFWAIPNQMNDGLGGDDMEMEMREIRKRLQLRNC LRILMGLSNHHHDHDEFLMP"

BASE COUNT 178 a 187 c 203 g 132 t
 ORIGIN

alignment_scores:
 Quality: 145.00 Length: 129
 Ratio: 2.014 Gaps: 3
 Percent Similarity: 55.814 Percent Identity: 28.682

alignment_block:

US-09-327-750D-34 x AF187066

Align seg 1/1 to: AF187066 from: 1 to: 700

18 LysLysAsnLysLysGlyGlyLysAlaSerLysGlnSerGluGluGlucose 34

```

||||:||||
162 AACAAAATCTCATCGGCAATGTCCACGAGAAAACCAAGAGATGGA 211
34 rHisHisLeuGluValGluAsnLysLysProGlyGly..... 47
212 GCAGCCCTGCAGAAATGGACAGGAAGACCCCTGTGGGAGGAGGTGAGG 261
47 ..... 47

262 GCCCAGCCCTGCTGCAAAACAACAACAACCAACCAACCAACCAACCAAC 311
48 .....AsnValArgArgLysValArgArgLeuValProAsnPheLeuTr 62
312 AACCAACCAAGAGAGAGCGGCTGCGGACTTGCCTTAACCTCCGATG 361
62 pAlaIleProAsnArgHisValAspHisSerGluGly.....GlyGluG 77
362 GGCATTTCCCAACAGCAGATGAATGACGGGTTGGGTGGAGATGAGATG 411
77 luValGlyArgPheValGlyGlnValMetGluAlaLysArgHisSerLys 93
412 ATATGGAATGTTCTACGACGAGATGAGAGATCCGGAGAAAGCTTAGG 461
94 GluGlnMetArgProTyrThrArgPheArgThrProGluProAspAs 110
462 GAGCTACAGCTGAGAAATGTTCTACGATCCTTATGGGGGAGCTGTCTAA 511
110 nHisTyrAsp.....PheCysLeuIlePro 118
512 CCACGAGATCACCAGATGAATTCGCCTTATGCCT 548

seq_name: gb_ro:AF097440
seq_documentation_block: 854 bp mRNA ROD 13-APR-1999
LOCUS AF097440 Mus musculus brain expressed x-linked protein 3 (Bex3) mRNA,
DEFINITION complete cds.
ACCESSION AF097440
VERSION AF097440.1 GI:4580593
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 854)
Brown,A.L. and Kay,G.F.
Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
Hum. Mol. Genet. 8 (4), 611-619 (1999)
99172070
Erratum:[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
2 (bases 1 to 854)
Brown,A.L. and Kay,G.F.
Direct Submission
Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
FEATURES
source
1. .854
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/db_xref="dbEST:AA272375"
/chromosome="X"
/map="near Plp"
/tissue_type="pooled organs"
1. .854
/gene="Bex3"
172. 546
/gene="Bex3"
/codon_start=1
/product="brain expressed X-linked protein 3"
/protein_id="AAD24431.1"

```

```

/db_xref="GI:4580594"
/translation="MANVHQQEEMEOPLONGQDRPVGGEGHQPAAANNHNNHNNH
HNHRRGOARLAPNFWAIPNQMNDCLGDDMEMFMEIRKRLRELQLRNC
LRIIMGELSNHHHDEFCLMP"
BASE COUNT 237 a 212 c 228 g 177 t
ORIGIN

alignment_scores:
Quality: 145.00 Length: 129
Ratio: 2.014 Gaps: 3
Percent Similarity: 55.814 Percent Identity: 28.682
alignment_block:
US-09-327-750D-34 x AF097440 ..
Align seg 1/1 to: AF097440 from: 1 to: 854
18 LysLysAsnLysLysGlyLysAlaSerLysGlnSerGluGluGlucose 34
157 AACAAATCTCATCGGCAATGTCCACGAGAAAACCAAGAGATGGA 206
34 rHisHisLeuGluValGluAsnLysLysProGlyGly..... 47
207 GCAGCCCTGCAGAAATGGACAGGAAGACCCCTGTGGGAGGAGGTGAGG 256
47 ..... 47
257 GCCCAGCGCTGCTGCAAAACAACAACAACCAACCAACCAACCAACCAAC 306
48 .....AsnValArgArgLysValArgArgLeuValProAsnPheLeuTr 62
307 AACCAACCAAGAGAGCGGCTGCGGACTTGCCTTAACCTCCGATG 356
62 pAlaIleProAsnArgHisValAspHisSerGluGly.....GlyGluG 77
357 GGCATTTCCCAACAGCAGATGAATGACGGGTTGGGTGGAGATGAGATG 406
77 luValGlyArgPheValGlyGlnValMetGluAlaLysArgHisSerLys 93
407 ATATGGAATGTTCTACGAGGAGATGAGAGATCCGGAGAAAGCTTAGG 456
94 GluGlnMetArgProTyrThrArgPheArgThrProGluProAspAs 110
457 GAGCTACAGCTGAGAAATGTTCTACGATCCTTATGGGGGAGCTGTCTAA 506
110 nHisTyrAsp.....PheCysLeuIlePro 118
507 CCACGAGATCACCAGATGAATTCGCCTTATGCCT 543

seq_name: gb_pr:HSV351F8
seq_documentation_block:
LOCUS HSV351F8 45678 bp DNA PRI 23-NOV-1999
DEFINITION Human DNA sequence from cosmid V351F8, between markers DXS366 and
DXS87 on chromosome X contains ESTs.
ACCESSION 270719
VERSION 270719.1 GI:1261915
KEYWORDS X.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 45678)
Whiteley,M.
Direct Submission
Submitted (09-APR-1995) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, UK. E-mail enquires: humquery@sanger.ac.uk
IMPORTANT: This sequence is the entire insert of clone V351F8. The
true left end of clone V351F8 is at 1 in this sequence. The true
right end of clone V351F8 is at 45678.
V351F8 is from the human chromosome X-specific cosmid library.
FEATURES
Location/Qualifiers

```

```

source
1. .45678
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="X"
/clone="GHC-351F8"
/clone_lib="SCcv"
1. .70
repeat_region
/partial
/note="Alu repeat: matches 80. .1 of consensus"
642. .930
repeat_region
/note="Alu repeat: matches 1. .308 of consensus"
1406. .5179
repeat_region
/note="L1 element fragment"
5243. .5276
repeat_region
/note="17 copies of 2 mer 82 % conserved"
5279. .5567
repeat_region
/partial
/note="Alu repeat: matches 304. .1 of consensus"
9975. .10369
repeat_region
/note="MSTA element fragment"
10399. .12017
repeat_region
/note="MSTAR element fragment"
11318. .11978
repeat_region
/note="THE1B element fragment"
11715. .11846
repeat_region
/note="MLTIR element fragment"
12074. .12170
repeat_region
/note="MSTA element fragment"
12227. .12486
repeat_region
/note="MSTA element fragment"
12437. .12502
repeat_region
/note="MSTC element fragment"
12756. .13059
repeat_region
/partial
/note="Alu repeat: matches 308. .1 of consensus"
14478. .14650
repeat_region
/note="THE1B element fragment"
14491. .14536
repeat_region
/note="MSTA element fragment"
14715. .14831
repeat_region
/note="THE1B element fragment"
15099. .15533
repeat_region
/note="L1 element fragment"
15534. .15626
repeat_region
/note="31 copies of 3 mer 85 % conserved"
15566. .15619
repeat_region
/note="3 copies of 18 mer 98 % conserved"
16352. .16436
repeat_region
/note="L1 element fragment"
16728. .16763
repeat_region
/note="9 copies of 4 mer 94 % conserved"
16764. .17054
repeat_region
/partial
/note="Alu repeat: matches 308. .1 of consensus"
19420. .19449
repeat_region
/note="15 copies of 2 mer 87 % conserved"
complement(19537. .20013)
misc_feature
/note="match: 3' EST N51315 clone 283089"
20071. .20114
repeat_region
/note="22 copies of 2 mer 98 % conserved"
20073. .20112
repeat_region
/note="10 copies of 4 mer 100 % conserved"
20073. .20126
repeat_region
/note="3 copies of 18 mer 87 % conserved"
23088. .23300
repeat_region
/note="L1 element fragment"
23311. .23490
repeat_region
/note="MLTIA element fragment"
23318. .23489
repeat_region
/note="MLTIB element fragment"
23362. .23491
repeat_region
/note="MLTIC element fragment"

repeat_region
23507. .24407
/note="L1 element fragment"
24376. .24462
repeat_region
/note="MSTC element fragment"
24379. .24491
repeat_region
/note="MSTA element fragment"
24688. .24738
repeat_region
/note="MSTA element fragment"
24810. .25016
repeat_region
/note="L1 element fragment"
25254. .25421
repeat_region
/note="L1 element fragment"
25439. .25696
repeat_region
/note="MER25 element fragment"
26284. .26343
repeat_region
/note="MLTID element fragment"
26428. .26491
repeat_region
/note="MLTID element fragment"
26465. .26518
repeat_region
/note="MLTIE element fragment"
27446. .27667
repeat_region
/note="L1 element fragment"
28954. .29247
repeat_region
/partial
/note="Alu repeat: matches 308. .1 of consensus"
30160. .30325
repeat_region
/note="2 copies of 83 mer 98 % conserved"
32588. .32875
misc_feature
/note="match: 5' EST H68599 clone 239077"
complement(32825. .33230)
misc_feature
/note="match: 3' EST H68239 clone 289077"
35551. .35802
repeat_region
/partial
/note="Alu repeat: matches 1. .260 of consensus"
35804. .35839
repeat_region
/note="18 copies of 2 mer 83 % conserved"
38625. .38705
repeat_region
/note="MLTIR element fragment"
38747. .38839
repeat_region
/note="MLTIR element fragment"
38949. .39032
repeat_region
/note="MLTIE element fragment"
38965. .39076
repeat_region
/note="MLTID element fragment"

BASE COUNT 14260 a 9135 c 9327 g 12956 t
ORIGIN

alignment_scores:
Quality: 138.00 Length: 127
Ratio: 2.000 Gaps: 3
Percent Similarity: 54.331 Percent Identity: 29.921

alignment_block:
US-09-327-750D-34 x HSV351F8 ..
Align seg 1/1 to: HSV351F8 from: 1 to: 45678

12 LeuThrValGluLysAspLysLysAsnLysGlyGlyLysAlaSerLy 28
||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
19408 TTATCTTCTAGGAAGAAAAAAGAAAAAAGAAAAAAGAAAAAATCAACAT 19457

28 sGlnSer.....GlUG 32
||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
19458 GGAAATGTCCCAAGAAAAACAAGTTGTGGAGAGCCCAAGTCGAGA 19507

32 luGluSerHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsn 48
||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
19508 ATGAAGCCCGCGCTTTAGGAGGTGGTGAATACAGGAGCCCTGGAGGAAT 19557

49 ValArgArgLysValArgArgLeuValProAsnPheLeuTrpAlaIlePr 65
||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
19558 GTTAAAGGGGTTTGGGGCTCCACCTGCCCGCGGGTTTGGAGAGGATGTGCC 19607

```

65 oAsnArgHisValAspHisSerGlu.....GlyGlyGluGluV 78
 ||||| ||||| : :
 19608 CAATAGGCTTCGCATACATGATGATGAGATGGAGATGATA 19657

78 alGlyArgPheValGlyGlnValMetGluAlaLysArgHisSerLysGlu 94
 : : ||||| : : : : || : : : :
 19658 TGGACGGCTTCAGGAGGATGAGAGCTAAGGAGAAATTAGGAA 19707

95 GlnGlnMetArgProTyrThrArgPheArgThrProGluPro..... 108
 ||||| : : : : : : : :
 19708 CTTGAGTGGAGTACAGTCTCGCATCTTTATAGGGACCTCTCTACCA 19757

109 .AspAsnHisTyrAspPheCysLeuIlePro 118
 ||||| : : : : : : : :
 19758 TGATCATCATGATGAGTTTGCCCTATGCGCT 19788

seq_name: gb_sts:G35294

seq_documentation_block: 477 bp DNA STS 02-OCT-1997
 LOCUS G35294 human STS SHGC-37409, sequence tagged site.
 DEFINITION G35294

ACCESSION G35294.1 GI:2459462
 VERSION STS; STS sequence; primer; sequence tagged site.
 KEYWORDS human.
 SOURCE

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
 Myers, R.M.
 JOURNAL
 COMMENT

Contact: Richard M. Myers
 Stanford Human Genome Center (SHGC)
 Stanford University School of Medicine
 Department of Genetics, M-344, Stanford, CA 94305, USA
 Tel: 4157259687
 Fax: 4157259689
 Email: myers@shgc.stanford.edu

Primer A: AACATCTTCCATGAAAGTTGATG
 Primer B: CTTTGGCATCTTCTGCAG
 STS size: 106
 PCR profile:

Initial incubation: 95 degrees C for 10 minutes

Denaturation: 94 degrees C for 15 seconds
 Annealing: 60 degrees C for 30 seconds
 Polymerization: 72 degrees C for 23 seconds
 PCR Cycles: 30
 Thermal Cycler: Perkin Elmer 9600

Protocol:
 Template: 25 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 AmpliTaq Gold Polymerase: 0.07 units/uL
 Total Vol: 5 uL

Buffer:
 MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from N51315
 -- Washington University/Merck EST sequence.

FEATURES

source
 1..477
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

STS
 primer_bind
 44..149
 44..167

primer_bind 153 a 115 c 79 g 130 t
 BASE COUNT
 ORIGIN

alignment_scores:
 Quality: 118.50 Length: 83
 Ratio: 2.370 Gaps: 2
 Percent Similarity: 60.241 Percent Identity: 33.735

alignment_block:

US-09-327-750D-34 x G35294/rev ..

Align seg 1/1 to reverse of: G35294 from: 1 to: 477

43 LysLysProGlyGlyAsnValArgLysValArgLysValArgLysValProAs 59

474 CAGGAGCTGGAGAAATGTTAAAGGGTTGGGCTCCACCTGCCCGG 425

59 nPheLeuTrpAlaIleProAsnArgHisValAspHisSerGlu..... 73

424 TTTTGGAGAGGATGTGCCCAATAGGCTTGCATAACATTTGATATGATAG 375

74GlyGlyGluGluValGlyArgPheValGlyGlnValMetGluAla 88

374 ATGGAGATGGAGATGATATGCAACGGTTTCATGGAGGAGATGAGAGACTA 325

89 LysArgHisSerLysGluGlnGlnMetArgProTyrThrArgPheArgTh 105

324 AGGAGGAAAATTAGGAACTTCAGTTGAGGTACAGTCTGGCGCATCTTAT 275

105 rProGluPro.....AspAsnHisTyrAspPheCysLeuIlePro 118

274 AGGGAGCCCTCCTCACCATCATCATGATGATGATGATGATGATGATGAT 226

seq_name: gb_ro:AF097439

seq_documentation_block:

LOCUS AF097439 785 bp mRNA ROD 13-APR-1999
 DEFINITION Mus musculus brain expressed X-linked protein 2 (Bex2) mRNA,
 complete cds.

ACCESSION AF097439
 VERSION AF097439.1 GI:4580591

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 785)

AUTHORS Brown, A.L. and Kay, G.F.

TITLE Bex1, a gene with increased expression in parthenogenetic embryos,

is a member of a novel gene family on the mouse X chromosome

JOURNAL Hum. Mol. Genet. 8 (4), 611-619 (1999)

MEDLINE 99172070

REMARK Erratum: [[published erratum appears in Hum Mol Genet 1999

May; 8(5):943]]

REFERENCE 2 (bases 1 to 785)

AUTHORS Brown, A.L. and Kay, G.F.

TITLE Direct Submission

JOURNAL Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of

Medical Research, Herston Rd, Brisbane, Qld 4029, Australia

FEATURES Location/Qualifiers

source

1..785

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/db_xref="dbEST:W48832"

/chromosome="X"

/map="near Plp"

/dev_stage="embryo; 15.5 dpc"

1..785

/gene="Bex2"

139..528

gene

CDS


```

alignment_scores:
  Quality: 112.50      Length: 128
  Ratio: 1.424        Gaps: 5
  Percent Similarity: 61.719  Percent Identity: 28.906

alignment_block:
  US-09-327-750D-34 x AF220189 ..

  Align seg 1/1 to: AF220189 from: 1 to: 828

      1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLys.. 16
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
167 ATGGAGTCCAAAGAGAACAAGCAGCAGTAAACAGTCTCAGCATGTGGAATGC 216
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

17 AspLysLysAsnLysGlyGlyLysAlaSerLysGlnSerGluGluG 33
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
217 CAACCAAGAAATGAGAGAAAGGACGAACTTGCTAATAAAGGGGAGCCCT 266
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
267 TGGCCCTCCCTTGGAGTCTCGTGAATACTGTGCTCAGAGGAATCGT 316
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
50 ArgArg...LysValArgArgLeuValProAsnPheLeuTrpAlaIlePr 65
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
317 AGGCGGTTCGCGTTAGGCACCCATCCTGCATATAGATGGCATATGAT 366
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
65 oAsnArgHisValAsp.....HisSerGluGlyGlyGluG 77
      : : : ||| : : : : : : : : : : : : : : : : : : : : :
367 GCATAGGCTTGGAGAACCCACAGCAGGATGAGAGAGAATATGGAAG 416
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
77 luValGlyArgPheValIcylGlnValMetGluAlaLysArgHisSerLys 93
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
417 CGATTGGGGAGGAGTGAGACAGCTGATGAAAGCTGAGG..... 457
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
94 GluGlnGlnMetArgProTyrThrArgPheArgThrProGluPro.... 108
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
458 GAAAGCAGGTGTAGTCATAGTCTCGGGGAGCTCAGCACTGACCCCTCA 507
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
109 ....AspAsnHisTyrAspPheCysLeuIlePro 118
      ||| : : ||| : : ||| ||| ||| ||| ||| ||| ||| ||| |||

```


OM of: US-09-327-750D-33 to: EST:* out_format : pfs

Date: Mar 11, 2002 2:16 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09327750/runat_l1032002_101153_20308/app_query.fasta_1.1472
-DB=EST -OPMT=fastap -SUFFIX=txt -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -OGAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09327750_8CGNI_L15654
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-33

Query length: 128

Database: EST:*

Database sequences: 11351937

Database length: 1077921985

Search time (sec): 4085.940000

score_list:

Sequence	Strd	Orig	zScore	EScore	Len	Documentation
gb_est1:AL022932	+	646.00	1047.10	3.7e-49	551	AL022932 v8130b28 Beddington mc
gb_est1:AW536404	-	646.00	1045.88	4.3e-49	642	AW536404 G0104A11-3 NIA Mouse E
gb_est2:BG064920	-	643.00	1040.25	8.9e-48	707	BG064920 H3025D01-3 NIA Mouse I
gb_est2:BG277659	+	633.00	1025.69	5.7e-48	575	BG277659 ux47c11-y1 Soares.NMMA
gb_est1:BG870503	+	633.00	1023.51	7.6e-48	756	BG870503 602791422F1 NCI CGAP_S
gb_est1:BE654459	+	615.00	997.65	2.1e-46	499	BE654459 UI-M-AJ1-aha-f-10-0-UI
gb_est1:AW536974	-	599.00	970.56	6.8e-45	577	AW536974 G0110H10-3 NIA Mouse E
gb_est2:BF607762	-	554.00	895.17	1.1e-40	785	BF607762 MYL_000704 Mouse 9-day
gb_est2:BG228077	-	553.00	895.87	9.8e-41	587	BG228077 ux47c11-x1 Soares.NMMA
gb_est2:BF608209	+	516.00	835.09	2.4e-37	650	BF608209 MYL_001088 Mouse 9-day
gb_est1:BE291071	+	505.50	788.18	1.9e-36	583	BE291071 G0108G311F1 NCI CGAP_M
gb_gss:AZ936393	-	493.50	798.78	2.5e-35	637	AZ936393 2M0193L05F Mouse 10kb
gb_est2:BG669326	-	489.50	795.78	3.7e-35	412	BG669326 DNAAEB11 Rat DRG Libr
gb_est2:BI183527	-	480.50	776.95	4.1e-34	701	BI183527 UNL-P-FN-by-f-07-0-UNL
gb_est1:AW536634	-	472.00	765.12	1.9e-33	549	AW536634 G0106H08-3 NIA Mouse E
gb_est2:BG655001	+	471.50	764.34	2.1e-33	547	BG655001 ib44f06-y1 HR85 islet
gb_est2:BG472167	+	471.50	763.17	2.4e-33	634	BG472167 602513894F1 NIH_MGC_16
gb_est2:BF971303	+	471.50	762.87	2.5e-33	658	BF971303 602273150F1 NIH_MGC_84
gb_est2:BG714823	+	471.50	762.68	2.6e-33	674	BG714823 602677146F1 NIH_MGC_96
gb_est1:AV702285	+	471.50	762.49	2.6e-33	690	AV702285 AV703285 ADB Homo sapi
gb_est1:AV705808	+	471.50	762.48	2.6e-33	691	AV705808 AV703808 ADB Homo sapi
gb_est1:AL526209	-	471.50	762.43	2.6e-33	695	AL526209 AL526209 LTI_NFL003_NE
gb_est2:BG820179	-	471.50	762.39	2.7e-33	699	BG820179 602782222F1 NCI CGAP_F
gb_est1:BE790774	+	471.50	761.48	3.0e-33	783	BE790774 G01581956F1 NIH_MGC_7
gb_est1:AL526247	+	471.50	761.37	3.0e-33	794	AL526247 AL526247 LTI_NFL003_NE
gb_est1:AV702643	+	468.50	757.71	4.8e-33	683	AV702643 AV702643 ADB Homo sapi
gb_est1:BE314909	+	468.00	756.01	4.7e-33	594	BE314909 G01140074F1 NIH_MGC_9
gb_est1:AL520931	-	468.00	756.51	5.7e-33	717	AL520931 AL520931 LTI_NFL004_NE
gb_est2:BE889313	+	468.00	755.92	6.1e-33	772	BE889313 G01513201F1 NIH_MGC_71
gb_est2:BG709503	+	468.00	755.54	6.4e-33	810	BG709503 602674781F1 NIH_MGC_96
gb_est2:BG707398	+	468.00	755.29	6.6e-33	835	BG707398 602672810F1 NIH_MGC_96
gb_est2:BG285825	+	468.00	754.30	7.5e-33	946	BG285825 602360917F1 NIH_MGC_93
gb_est2:BF038787	+	468.00	753.53	8.3e-33	1042	BF038787 G01462119F1 NIH_MGC_6
gb_est1:AI929703	+	464.50	752.59	9.4e-33	576	AI929703 au63f04-y1 Schneider f
gb_est2:BG715659	+	464.50	750.65	1.2e-32	734	BG715659 602676924F1 NIH_MGC_96
gb_est2:BG714974	-	461.00	745.07	2.5e-32	726	BG714974 602677263F1 NIH_MGC_96
gb_est2:BF967675	-	461.00	744.36	2.7e-32	793	BF967675 602287383F1 NIH_MGC_96
gb_est2:BG707734	+	461.00	744.12	2.8e-32	817	BG707734 602671235F1 NIH_MGC_96
gb_est2:BG705843	+	457.00	737.68	6.3e-32	813	BG705843 602669329F1 NIH_MGC_96
gb_est2:BG709325	+	456.50	736.72	7.2e-32	828	BG709325 602673345F1 NIH_MGC_96

gb_est1:AI291270 + 455.00 738.14 6.0e-32 511
gb_est1:AL024066 + 453.00 734.67 9.3e-32 526
gb_est1:BE266012 + 453.00 733.49 1.1e-31 610
gb_est1:AI291126 + 453.00 732.75 1.2e-31 669
gb_est2:BF237433 + 453.00 732.27 1.3e-31 711
seq_name: gb_est1:AL022932

seq_documentation_block:
LOCUS AL022932 551 bp mRNA EST
DEFINITION v8130b28 Beddington mouse dissected endoderm Mus musculus CDNA
clone 528-3B10 5', mRNA sequence.
ACCESSION AL022932
VERSION AL022932.1 GI:6645505
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 551)
Harrison,S.M., Dunwoodie,S.L., Arkell,R.M., Lehrach,H. and
Beddington,R.S.

Isolation of novel tissue-specific genes from cDNA libraries
representing the individual tissue constituents of the gastrulating
mouse embryo
Development 121 (8), 2479-2489 (1995)
95401865

CONTACT: Wiles,M., Lehrach,H. and Ayner,P.
ECC Mouse Transcript Mapping Consortium
Genoscope - CNS
2, rue Gaston Cremieux, 91000 Evry, France
Email: pavner@pasteur.fr

Clone available from Ressourcenzentrum, Heubnerweg 6, D-14059
Berlin, Germany. Web site http://www.rzpd.de
Seq primer: CCGTCCGAATCCGGGT;
High quality sequence only submitted.
Vector: pSport1; site_1: NotI; site_2: SalI;
Cloned unidirectionally.
Dissected endoderm 7.5 days.

Average insert size: 1.2 kb (range: 0.2 - 2.kb).
Location/Qualifiers
1..551

/organism="Mus musculus"
/strain="C57Bl6 x DBA"
/db_xref="taxon:10090"
/clone="528-3B10"
/clone_lib="Beddington mouse dissected endoderm"
/tissue_type="dissected endoderm"
/dev_stage="7.5 dpc"

/note="Vector: pSport1; Site_1: NotI; Site_2: SalI; Cloned
unidirectionally ~ High quality sequence only submitted. ~
Average insert size: 1.2 kb (range: 0.2 - 2.kb)"

BASE COUNT 153 a 124 c 191 g 83 t

alignment_scores:
Quality: 646.00 Length: 128
Ratio: 5.168 Gaps: 0
Percent Similarity: 97.656 Percent Identity: 92.969

alignment_block:
US-09-327-750D-33 x AL022932

Align seg 1/1 to: AL022932 from: 1 to: 551

1 MetGlySerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspH 17
|||||

142 ATGGAGTCCAAGATCAAGCGCTGAAATATCAACATGAGATGACCA 191
|||||

17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleLysArg 34
|||||

192 TCAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 241
|||||

```

34 luProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
|||||:||||| ||| |||||||||||||||||||
242 AGCCAGCTGTGGCCCTCGACTCCGAGGCTGGCAAAACTGTGCGCTAGA 291
|||||:||||| ||| |||||||||||||||||||
51 GlyGlyArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67
|||||:||||| ||| |||||||||||||||||||
292 GGAGGTGCGAGGCGGTTCGGGTTCGCGAGCCCATCGCTCACTATAGATG 341
|||||:||||| ||| |||||||||||||||||||
67 pAspLeuMethHisArgValGlyGluProGlnGlyArgMetArgGluGlu 84
|||||:||||| ||| |||||||||||||||||||
342 GGACCTGATCGACAGGCTTGGGAGGCCCGGAGGAGGATGAGAGAGGAGA 391
|||||:||||| ||| |||||||||||||||||||
84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
|||||:||||| ||| |||||||||||||||||||
392 ACGTACAGAGGTTGGGGGTGTGTGAGACAGCTCATGGAGAAGCTGAGG 441
|||||:||||| ||| |||||||||||||||||||
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||:||||| ||| |||||||||||||||||||
442 GAAGGCAGCTGAGCCACAGCCTGCGGCGGTGTAGCACTGACCGGCTCA 491
|||||:||||| ||| |||||||||||||||||||
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||:||||| ||| |||||||||||||||||||
492 TCATGACCACCATGATGAGTTTGGCTCATGCC 525
|||||:||||| ||| |||||||||||||||||||

```

seq_name: gb_est1:AW536404

```

seq_documentation_block:
LOCUS AW536404 642 bp mRNA EST 31-AUG-2000
DEFINITION G0104A11-3 NTA Mouse E7.5 Embryonic Portion cDNA Library Mus
musculus cDNA clone G0104A11 3', mRNA sequence.
ACCESSION AW536404
VERSION AW536404.1 GI:7178821
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac
,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H.
III., Becker,K.G. and Ko,M.S.H.
TITLE Genome-wide expression profiling of mid-gestation placenta and
embryo using a 15,000 mouse developmental cDNA microarray
Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
20381348
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: G0104 row: A column: 11
Seq primer: -21M13 Forward
High quality sequence stop: 642
POLYA=Yes.

```

```

FEATURES
    source
        1..642
            /organism="Mus musculus"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /clone="G0104A11"
            /clone_lib="NTA Mouse E7.5 Embryonic Portion cDNA Library"
            /sex="unknown"
            /dev_stage="7.5dpc Embryo"
            /lab_host="DH10B"
            /note="Vector: pSPOR1 (Gibco/BRL Life Technology);
Site.1: SalI; Site.2: NotI; Total RNAs were extracted from
6 Embryo. The double-stranded cDNA was synthesized by
Gibco's kit with an Oligo(dT) primer [NotI primer-adapter
from GibcoBRL]
[5'-pGACTAGTCTTATGATCGAGCGCGCCCTTTTCTTTTCTTTT-3']
from 0.5lug of mRNA . The double-stranded cDNAs were
treated with T4 DNA polymerase and purified by

```

ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal3 (include SalI sequence). The cDNAs were purified by phenol/chloroform and separated from free linkers by Centricon 100. Then, cDNAs were amplified by long-range high fidelity PCR using Takara's Ex Taq polymerase. Then, the cDNAs were purified by phenol/chloroform and by Centricon 100. The cDNAs were digested with SalI and NotI enzymes. Then, the cDNAs were size selected by Gibco's Size Fractionation Column. The cDNAs were cloned into SalI/NotI site of pSPOR1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by chemical method. The library was constructed by Xiaohong Wang and Minoru S. H. Ko."

BASE COUNT 145 a 182 c 133 g 182 t
ORIGIN

alignment_scores:
Quality: 646.00 Length: 128
Ratio: 5.168 Gaps: 0
Percent Similarity: 97.656 Percent Identity: 92.969

alignment_block:
US-09-327-750D-33 x AW536404/rev ..

Align seg 1/1 to reverse of: AW536404 from: 1 to: 642

```

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHi 17
|||||:||||| ||| |||||||||||||||||||
606 ATGGAGTCCAAAGATCAAGGCGTGAATAAATCTCAACATGGAGAATGACCA 557
|||||:||||| ||| |||||||||||||||||||
17 sGlnLysLysGluGluLysGluGlnAspThrIleLysArgG 34
|||||:||||| ||| |||||||||||||||||||
556 TCAGAAAAAGAGGAGGAGGAAGAAACCAAGATACCATCAGAAAGG 507
|||||:||||| ||| |||||||||||||||||||
34 luProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
|||||:||||| ||| |||||||||||||||||||
506 ACCGAGCTGTGGCCCTCGACTCCGAGGCTGGCAAAACTGTGCACCTAGA 457
|||||:||||| ||| |||||||||||||||||||
51 GlyGlyArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67
|||||:||||| ||| |||||||||||||||||||
456 GGAGGTGCGAGGCGGTTCGGGTTCGGCAGCCCATCGCTCACTATAGATG 407
|||||:||||| ||| |||||||||||||||||||
67 pAspLeuMethHisArgValGlyGluProGlnGlyArgMetArgGluGlu 84
|||||:||||| ||| |||||||||||||||||||
406 GGACCTGATCGACAGGCTTGGGAGGCCCGGAGGATGAGAGAGGAGA 357
|||||:||||| ||| |||||||||||||||||||
84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
|||||:||||| ||| |||||||||||||||||||
356 ACGTACAGAGGTTGGGGGTGTGTGAGACAGCTCATGGAGAAGCTGAGG 307
|||||:||||| ||| |||||||||||||||||||
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||:||||| ||| |||||||||||||||||||
306 GAAAGGCAGCTGAGCCACAGCCTGCGGCGGTGTAGCACTGACCGGCTCA 257
|||||:||||| ||| |||||||||||||||||||
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||:||||| ||| |||||||||||||||||||
256 TCATGACCACCATGATGAGTTTGGCTCATGCC 223
|||||:||||| ||| |||||||||||||||||||

```

seq_name: gb_est2:BG064920

```

seq_documentation_block:
LOCUS BG064920 707 bp mRNA EST 26-JAN-2001
DEFINITION H3025D01-3 NTA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
H3025D01 3', mRNA sequence.
ACCESSION BG064920
VERSION BG064920.1 GI:12547483
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 707)

```

AUTHORS Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
TITLE T.S., Carter,M.G. and Ko,M.S.H.
JOURNAL Verification and initial annotation of NIA mouse 15K cDNA clone set
COMMENT Unpublished (2001)
 Other_ESTs: H3025D01-5
 Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please
 visit <http://igsun.grc.nia.nih.gov/cDNA/15k.html> for details.
 Plate: H3025 row: D column: 01
 Seq primer: -21M13 Forward
 High quality sequence stop: 707
 POLYA=Yes.

FEATURES source

Location/Qualifiers

1. 707

```
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="H3025D01"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/notes=Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearrayed set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos , and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the x chromosome, 1998, Hum Mol Genet 7: 1967-1978."
```

a 209 c 146 g 198 t

```
alignment_scores:      Length: 128
                       Quality: 643.00
                       Ratio: 5.114
                       Gaps: 0
                       Percent Identity: 92.188
                       Percent Similarity: 97.656
```

alignment block:

US-09-327-750D-33 x BG064920/rev

Align seq 1/1 to reverse of: BG064920 from: 1 to: 707

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHis 17
1606 ATGGAGTCCAAAGATCAAGCGTGAATAAATCTCAACATGGAGAAATGACCA 557
17 sGlnLysLysGluLysGlnGlyLysProGlnAspThrIleLysArgG 34
1556 TCAGAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 507
34 LuProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
1506 AGCAGCTGTGGCCCTCACCCTCCGAGGCTGGCAAAACTGTGCAACCTAGA 457
51 GlyGlyArgArgArgPheArgValArgGlnProIleSerHisThrArgTr 67

```
US-09-327-750D-33 x BG277659
Align seg 1/1 to: BG277659 from: 1 to: 575

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuMetGluAsnAspHI 17
192 ATGGAGTCCAAAGATCAAGCGGTCAAAAATCTCAACATGGAGAATGACCA 241

17 sGlnLysLysGluGluLysGluGluLysProGlnAspThrIleLysArgG 34
242 TCAGAAAAGAGGAGAGGAAGAAAGAAAGCCACCAAGATACCATCAGAAGG 291

34 luProValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
292 AGCCAGCTGTGGCCTCACCTCCGAGCTGGCAAAACTGTGCACCTAGA 341

51 GlyGlyArgArgPheArgValArgGlnProIleSerHisTyArgTr 67
342 GGAGGTCCAGAGCGGTTCGGGTTCCGAGCCCATCGCTCACTATAGATG 391

67 pAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluGluA 84
392 GGACCTGATCAGAGGGTTGGGAGCCCCAGGGAAGGATGAGAGAGGAGA 441

84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
442 ACGTACAGAGTTGNGGGTGATGTGAGACAGCTCATGGAGAGCTGAGG 491

101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHI 117
492 GAAAGGCAGCTGAGCCACAGCCTCGCGGGCTTAGCAGCTACCCCGCTCA 541

117 sHisAspHisHisAspGluPheCysLeuMetPro 128
542 TCATGACCAACCATGATGAGTTGTGCCTCATGCC 575

seq_name: gb_est2:BG870503

seq_documentation_block:
LOCUS BG870503 756 bp mRNA EST 29-MAY-2001
DEFINITION 602791422F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4922644 5',
mRNA sequence.
ACCESSION BG870503
VERSION BG870503.1 GI:14221043
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10842 row: 1 column: 05
High quality sequence stop: 756.
Location/Qualifiers
1..756
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:

NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 207 a 164 c 230 g 155 t
ORIGIN

alignment_scores:
Quality: 633.00 Length: 129
Ratio: 5.064 Gaps: 1
Percent Similarity: 96.899 Percent Identity: 92.248

alignment_block:
US-09-327-750D-33 x BG870503
Align seg 1/1 to: BG870503 from: 1 to: 756

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuMetGluAsnAspHI 17
165 ATGGAGTCCAAAGATCAAGCGGTCAAAAATCTCAACATGGAGAATGACCA 214

17 sGlnLysLysGluGluLysGluGluLysProGlnAspThrIleLysArgG 34
215 TCAGAAAAGAGGAGAGGAAGAAAGAAAGCCACCAAGATACCATCAGAAGG 264

34 luProValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
265 AGCCAGCTGTGGCCTCACCTCCGAGCTGGCAAAACTGTGCACCTAGA 314

51 GlyGlyArgArgPheArgValArgGlnProIleSerHisTyArgTr 67
315 GGAGGTCCAGAGCGGTTCGGGTTCCGAGCCCATCGCTCACTATAGATG 364

67 pAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluGlu 83
365 GGACCTGATCAGAGGGTTGGGAGCCCCAGGGAAGGATGAGAGAGGAG 414

84 AsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
415 AACCTACAGAGGTTTGGGGTGATGTGAGACAGCTCATGGAGAAGCTGAG 464

100 gGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 117
465 GGAAGGCAGCTGAGCCACAGCCTCGCGGGCTTAGCAGCTACCCCGCTC 514

117 isHisAspHisHisAspGluPheCysLeuMetPro 128
515 ATCATGACCAACCATGATGAGTTTGCCTCATGCC 549

seq_name: gb_est1:BE654459

seq_documentation_block:
LOCUS BE654459 499 bp mRNA EST 06-SEP-2000
DEFINITION UI-M-AJ1-aha-f-10-0-UI.r1 NIH_BMAP_MOB_N Mus musculus cDNA clone
UI-M-AJ1-aha-f-10-0-UI 5', mRNA sequence.
ACCESSION BE654459
VERSION BE654459
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Bonardo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL 9704477
MEDLINE
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
```

Fax: 301 443 9890

Email: mEST@mail.nih.gov

cDNA Library Preparation: M.B. Soares Lab Clone Distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements

Seq primer: M13 Reverse.

Location/Qualifiers

1. .499

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-AJ1-aha-f-10-0-UI"

/clone_lib="NIH_BMAP_MOB_N"

/dev_stage="27-32 days"

/lab_host="DHI08 (Life Technologies)"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; The

NIH_BMAP_MOB_N library is a normalized library constructed

from mouse olfactory bulbs. The tag is a string of 5

nucleotides present between the Not I site and the

oligo-dT track. The library was constructed as described

by Bonaldo, Lennon and Soares, Genome Research 6: 791-806

, 1996. Tissue provided by Ms. Annie Novakovich,

Zivic-Miller Laboratories."

146 a 108 c 179 g 56 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 615.00 Length: 123

Ratio: 5.125 Gaps: 0

Percent Similarity: 97.561 Percent Identity: 92.693

alignment_block:

US-09-327-750D-33 x BE654459 ..

Align seg 1/1 to: BE654459 from: 1 to: 499

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHi 17

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

131 ATGGAGTCCAAAGATCAAGCGCTGAAATCTCAACATGGAGATGACCA 180

17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleLysArgG 34

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

181 TCAGAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 230

34 luProValAlaAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

231 AGCCAGCTGTGCGCTGACCTCCGAGGCTGGCAAAACTGTCCACTAGA 280

51 GlyGlyArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

281 GGAGGTCGAGCGGTTCGGGTTCGGAGCCCATCGCTCACTATAGATG 330

67 pAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluGlu 84

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

331 GGACCTGATGACAGAGGTTGGGGAGGCCCGAGGAGGATGAGAGGAGA 380

.84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

381 ACGTACAGAGTTGGGGGTGATGTGACAGACGCTCATGAGAGCTGAGG 430

101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

431 GAAAGCAGCTGAGGCACACAGCTGGGGGGTGTAGCACTGACCCGCTCA 480

117 sHisAspHisHisAspGlu 123

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

481 TCATGACCACCATGATGAG 499

seq_name: gb_estl:AW536974

seq_documentation_block:

LOCUS AW536974 577 bp mRNA EST 31-AUG-2000

DEFINITION G0110H10-3 NIA Mouse E7.5 Embryonic Portion cDNA Library Mus

musculus cDNA clone G0110H10 3', mRNA sequence.

ACCESSION AW536974

VERSION AW536974.1 GI:7179391

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 577)

AUTHORS Tanaka, T.S., Jaradat, S.A., Lim, M.K., Kargul, G.J., Wang, X., Grahovac

M.J., Pantano, S., Sano, Y., Plao, Y., Nagaraja, R., Dol, H., Wood, W.H.

III, Becker, K.G. and Ko, M.S.H.

Genome-wide expression profiling of mid-gestation placenta and

embryo using a 15,000 mouse developmental cDNA microarray

Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)

20381348

Contact: George J. Kargul

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@lgsun.grc.nia.nih.gov

Plate: G0110 row: H column: 10

Seq primer: -21M13 Forward

High quality sequence stop: 577

POLYA-Yes.

FEATURES

source

1. .577

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="G0110H10"

/clone_lib="NIA Mouse E7.5 Embryonic Portion cDNA Library"

/sex="unknown"

/dev_stage="7-5dpc Embryo"

/lab_host="DHI08"

/note="Vector: pSPORT1 (Gibco/BRL Life Technology);

Site_1: SalI; Site_2: NotI; Total RNAs were extracted from

6 Embryo. The double-stranded cDNA was synthesized by

Gibco's kit with an Oligo(dT) primer (NotI primer-adaptor

from GibcoBRL)

[5'-pgACTAGTCTAGATCGGAGCGCCCTTTTCTTTTCTTTT-3']

from 0.5ug of mRNA. The double-stranded cDNAs were

treated with T4 DNA polymerase and purified by

ethanol-precipitation. The cDNAs were ligated to

lone-linker LL-Sal3 (include SalI sequence). The cDNAs

were purified by phenol/chloroform and separated from

free linkers by Centricon 100. Then, cDNAs were amplified

by long-range high fidelity PCR using Takara's Ex Taq

polymerase. Then, the cDNAs were purified by

phenol/chloroform and by Centricon 100. The cDNAs were

digested with SalI and NotI enzymes. Then, the cDNAs were

size selected by Gibco's Size Fractionation Column. The

cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid

vector. The DHI08 E. coli host was transformed with the

ligation mixture by chemical method. The library was

constructed by Xiaohong Wang and Minoru S. H. Ko."

BASE COUNT 137 a 160 c 126 g 154 t

ORIGIN

alignment_scores:

Quality: 599.00 Length: 118

Ratio: 5.209 Gaps: 0

Percent Similarity: 97.458 Percent Identity: 93.220

alignment_block:

US-09-327-750D-33 x AW536974/rev ..

OM of: US-09-327-750D-35 to: N_Geneseq_1101.* out_format : pfs
Date: Mar 11, 2002 3:42 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09327750/runat_11032002_101154_20362/app_query.fasta_1.1472
-DB=N_Geneseq_1101 -QFMT=fastap -SUFFIX=eng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsu62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US09327750_@CGN1_330 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-327-750D-35
Query length: 117
Database: N_Geneseq_1101.*
Database sequences: 930621
Database length: 428662619
Search time (sec): 355.560000

score_list:	Sequence	Strd	Orig	zscore	EScore	Len	Documentation
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517	+		355.50	767.81	1.1e-34	865
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13750	+		355.50	764.27	1.7e-34	1229
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH85548	+		355.50	763.22	2.0e-34	1364
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH35259	+		175.00	367.98	2.0e-12	891
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH21748	+		174.00	365.47	2.8e-12	917
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH23528	+		153.00	321.71	7.6e-10	700
	/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01004	+		147.00	312.12	2.6e-09	485
	/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01005	+		141.00	297.91	1.6e-08	532
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH75810	+		99.50	202.37	0.0034	767
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH58581	+		99.50	201.19	0.0039	862
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH59511	+		99.50	200.78	0.0042	898
	/SID22/gcgdata/geneseq/geneseq/NA199.DAT:AAH23519	+		98.50	159.28	0.8526	4453
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143	+		94.00	189.87	0.0169	792
	/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01003	+		92.00	192.29	0.0124	401
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH12021	+		89.00	189.41	0.0179	276
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH14501	-		89.00	189.41	0.0179	276
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH05905	-		89.00	189.41	0.0179	276
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13308	+		87.50	181.99	0.0463	415
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH134661	+		87.50	181.99	0.0463	415
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03209	+		87.50	181.99	0.0463	415
	/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH03880	+		87.50	177.29	0.0847	662
	/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH03879	+		87.50	176.75	0.0907	698
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH60367	+		87.50	174.68	0.1184	858
	/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH10889	+		84.50	170.20	0.2101	692
	/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH70191	+		82.00	157.98	1.01	1344
	/SID22/gcgdata/geneseq/geneseq/NA199.DAT:AAH06000	+		81.50	150.50	2.63	2533
	/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH46139	+		80.50	118.25	164.41	50000
	/SID22/gcgdata/geneseq/geneseq/NA198.DAT:AAH21511	+		80.00	137.30	14.29	6755
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH17145	+		79.50	149.29	3.07	1839
	/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH37111	+		77.50	141.89	7.93	2471
	/SID22/gcgdata/geneseq/geneseq/NA199.DAT:AAH13331	+		77.50	124.33	75.44	14141
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH29749	+		76.00	133.90	22.10	3929
	/SID22/gcgdata/geneseq/geneseq/NA199.DAT:AAH13264	+		76.00	129.63	38.21	6005
	/SID22/gcgdata/geneseq/geneseq/NA199.DAT:AAH31091	+		75.50	144.44	5.72	1236
	/SID22/gcgdata/geneseq/geneseq/NA198.DAT:AAH06613	+		75.50	129.01	41.37	5721
	/SID22/gcgdata/geneseq/geneseq/NA198.DAT:AAH93167	+		75.50	127.91	47.63	6381
	/SID22/gcgdata/geneseq/geneseq/NA2001.DAT:AAH28893	+		75.00	143.96	6.08	1161
	/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH38803	+		74.50	140.84	8.45	1342
	/SID22/gcgdata/geneseq/geneseq/NA2000.DAT:AAH49751	+		74.50	141.39	8.46	1343
	/SID22/gcgdata/geneseq/geneseq/NA197.DAT:AAH12500	+		74.50	127.88	47.82	5137

/SID2/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ01022 - 74.50 103.74 1.1e+03
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAI21493 + 74.00 152.63 2.00
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAI46784 + 74.00 152.63 2.00
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAI07190 + 74.00 152.63 2.00
/SID2/gcgdata/geneseq/geneseq/NA1995.DAT:AAQ086237 + 74.00 128.11 46.45

seq_name: /SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517

seq_documentation_block:

ID AAH03517 standard; cDNA: 865 BP.

AC AAH03517;

DT 26-JUN-2001 (first entry)

DE Human cDNA clone (5'-primer) SEQ ID NO:352.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

PN EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

PA Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

PS Claim 1; SEQ ID 352; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

XX Sequence 865 BP; 239 A; 177 C; 224 G; 221 T; 4 other;

alignment_scores:		
Quality:	355.50	Length: 120
Ratio:	3.665	Gaps: 3
Percent Similarity:	80.833	Percent Identity: 59.167

```

Align seg 1/1 to: AAH13750 from: 1 to: 1229

1 MetAlaSerLysLysGlnValIleLeu...AspLeuThrValGluLysAS 16
  ||| ||||| ::::: ::::: |||||: |||||:
191 ATGGAGTCCAAGAGGAACATAGCGGCAAAACAATCTCAACGGGGAAATGC 240
  ||| ||||| ::::: ::::: |||||: |||||:
216 pLysLysAspLysArgGlyGly...LysAlaSerLysGlnSerGluGluG 32
  ||| ||||| ::::: |||||: |||||: |||||: |||||: |||||:
241 CCAACAAGAAAACCAAGAGGGGAGCAGGGCCGCCACGCAAGATGACAGAG 290
  ||| ||||| |||||: |||||: |||||: |||||: |||||: |||||:
32 LuProHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnVal 48
  || ||||| |||||: |||||: |||||: |||||: |||||: |||||:
291 AATCCCGCATTTCCGAGGGGGTGAAGCCGACAGAGCCTGGAGGAATATC 340
  ||||| ||||| ::::: |||||: |||||: |||||: |||||: |||||:
49 ArgArg...LysValArgArgLeuValProAsnPheLeuTrpAlaIlePr 64
  ||||| ::||| |||||: |||||: |||||: |||||: |||||: |||||:
341 AGCGGGGGCGAGTTAGCGCACTGTCTCCCTAATTTTCGATGGGCCATACC 390
  ||||| ||||| ::::: |||||: |||||: |||||: |||||: |||||:
64 oAsnArgHisValAspArgAsnGluGlyGlyGluAspValGlyArgPheV 81
  ||||| ||||| ::::: |||||: |||||: |||||: |||||: |||||:
391 TAATAGGCATATTGAGGCACAATGAAGCGACAGATGATGTAGAAAGGTTTG 440
  ||||| ||||| ::::: |||||: |||||: |||||: |||||: |||||:
81 alValGlnGlyThrGluValLysArgLysThrThrGluGlnGlnValArg 97
  ||||| ||||| ::::: |||||: |||||: |||||: |||||: |||||:
441 TAGGCAGATGTCGAAATCAAGAGAAGAAGACTAGGGAACACGACAGATGAGG 490
  ||||| ||||| ::::: |||||: |||||: |||||: |||||: |||||:
98 ProTyrArgArgPheArgThrProGluProAsnHisTyrAspPheCy 114
  ||||| ||||| ::::: |||||: |||||: |||||: |||||: |||||:
491 CACTATATGGCTTTCCAAACCTCTGAACCTGCACCAACCATATGACATTTTG 540
  ||||| ||||| ::::: |||||: |||||: |||||: |||||: |||||:

```

```

seq_documentation_block:
ID   AAHL13750 standard; cDNA; 1229 BP.
XX
XX   AAHL13750;
XX
XX   26-JUN-2001 (first entry)
DT
XX
XX   Human cDNA sequence SEQ ID NO:10656.
DE
XX
XX   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KW
XX   Homo sapiens.
XX
XX   EP1074617-A2.
PN
XX
XX   07-FEB-2001.
PD
XX
XX   28-JUL-2000; 2000EP-0116126.
PF
XX
XX   29-JUL-1999; 99JP-0248036.
PR
XX   27-AUG-1999; 99JP-0300253.
PR
XX   11-JAN-2000; 2000JP-0118776.
PR
XX   02-MAY-2000; 2000JP-0183767.
PR
XX   09-JUN-2000; 2000JP-0241899.
XX
XX   (HELI-) HELIX RES INST.
PA
XX
XX   Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI   Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

```


CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 485 BP; 125 A; 113 C; 159 G; 84 T; 4 other;

alignment_scores:
Quality: 147.00 Length: 87
Ratio: 2.625 Gaps: 1
Percent Similarity: 64.368 Percent Identity: 39.080
alignment_block:
US-09-327-750D-35 x AAC01004 ..
Align seg 1/1 to: AAC01004 from: 1 to: 485

28 GlnSerGluGluGluProHisHisLeuGluGluValGluAsnLysLysPr 44
||||: ||||| :||| ||||| :|||
217 CAGAAATGGAGGAGGAAGACCGCCCTTTGGGAGGAGGTCAAGGCCACCGCC 266
44 oGlyGlyAsnValArgArgLysValArgArgLeuValProAsnPheLeuT 61
||||: ||||| :||| :|||: ||||| :|||: ||||| :|||: ||||| :|||
267 TGCAGGAATTCGACGGGGACAGCGCTCGCGACTGCGCCCTAAATTTTCGAT 316
61 rPalatleProAsnArgHisValAspArgAsnGluGly.....GlyGlu 75
||||: ||||| :||| :|||: ||||| :|||: ||||| :|||: ||||| :|||
317 GGCCATACCATAGGCAGATCAATGATGGGATGGGTGATGATGGAGAT 366
76 AspValGlyArgPheValValGlnGlyThrGluValLysArgLysThrTh 92
||||: ||||| :||| :|||: ||||| :|||: ||||| :|||: ||||| :|||
367 GATATGGAATATTCATGGAGGAGATGAGAGAAATCAGAAGAAACTTAG 416
92 rGluGlnGlnValArgProTyrArgArgPheArgThrProGluProAspA 109
||||: ||||| :||| :|||: ||||| :|||: ||||| :|||: ||||| :|||
417 GGAGTGCACTCAGGAATGTCTGCGTATCCTTATGGGGGAKCTCTCTA 466
109 snHistyrAsp 112
||||: ||||| :||| :|||: ||||| :|||: ||||| :|||: ||||| :|||
467 ATCACCATGAC 477

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAC01005

seq_documentation_block:
ID AAC01005 standard; cDNA; 532 BP.
XX
AC AAC01005;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 1003.
XX
DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX (GBST) GENSET.
PA

PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
DR P-PSDB; AAG00999.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 1003; 71pp + CD-ROM; English.

CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 532 BP; 151 A; 121 C; 163 G; 91 T; 6 other;

alignment_scores:
Quality: 141.00 Length: 112
Ratio: 2.074 Gaps: 3
Percent Similarity: 60.714 Percent Identity: 33.036
alignment_block:
US-09-327-750D-35 x AAC01005 ..
Align seg 1/1 to: AAC01005 from: 1 to: 532

4 LysLysGlnValIleLeu...AspLeuThrValGluLysAspLysLysAs 19
||||: ||||| :||| :|||: ||||| :|||: ||||| :|||: ||||| :|||
204 AAAAAAATCTCATCTGCAATATTTACACGAGAAACGAAAGACAGATGGA 253
19 pLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluProHisHisL 36
||||: ||||| :||| :|||: ||||| :|||: ||||| :|||: ||||| :|||
254 GCAG.....CCTATGCAGAAATGGAGAGGAGACCGCCCTT 288
36 euGluGluValGluAsnLysLysProGlyGlyAsnValArgArgLysVal 52
|| ||||: ||||| :|||: ||||| :|||: ||||| :|||: ||||| :|||
289 TGGGAGGAGGTCAAGCCACCAGCCTGCAGAAATCGACGGGAMAGGCT 338
53 ArgArgLeuValProAsnPheLeuTrpAlaIleProAsnArgHisValAs 69
||||: ||||| :||| :|||: ||||| :|||: ||||| :|||: ||||| :|||
339 CGCCGAYTTGCCCTAAATTTTCGATGGCCCATACCCAATAGCAGATCAA 388
69 pArgAsnGluGly.....GlyGluAspValGlyArgPheValValGlnG 84
||||: ||||| :||| :|||: ||||| :|||: ||||| :|||: ||||| :|||
389 TGATGGATGGGTGAGATGGAGATCATATGGAATATTCATGGAGGAGA 438
84 lyThrGluValLysArgLysThrThrGluGlnGlnValArgProTyrArg 100
||||: ||||| :||| :|||: ||||| :|||: ||||| :|||: ||||| :|||
439 TCAGAGAAATCAGAGAAACCTTAGGAGCTGCAKTTGAGGAATTCCTCG 488
101 ArgPheArgThrProGluProAspAsnHisTyrAsp 112
||||: ||||| :||| :|||: ||||| :|||: ||||| :|||: ||||| :|||
489 CGTATCCTTATGGGGGAKCTCTCTAATCACCATGAC 524

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH75810
seq_documentation_block:
ID AAH75810 standard; cDNA; 767 BP.
XX
AC AAH75810;


```
Percent Similarity: 62.745 Percent Identity: 26.471
alignment_block:
US-09-327-750D-35 x AAX23519 ..
Align seg 1/1 to: AAX23519 from: 1 to: 44453

1 MetaLaSerLysLysGlnValIleLeuAspLeuThrValGluLysAspLy 17
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
43713 CTACAAAGAGAGAGACACTGGCTTTTCCAGCTTCAGCTTGAGGAAGAGAG 43762
```

17 sLysAspLysArqGlvGlvLysAlaSerLysGlnSerGluGluProH 34

43763 GGAGGAAAAAGGAGGGGAGGAAGAGCAAAAGGAGGAGGAGGAAA 43812

34 ISHISLEUGIUVAIGUASHNLYSYPGGLY:..GLYASHVATARGA 30
||||| :.....|:|:| :|:|:| :|:|:|

43813 TGGACACGGAGGAAAAAAGGAGGAGGATGAGGAGGGAAAAAGGA 43862

30 TGLYSVAIATGATGLeuvalTFOASHFneLEUTIPATAITTEPTOASIIAY 88
||| |:::|| ||| ::::: ||| :::::

67 HisValAspArgAsnGluGlyCysAspValGlyArgPheValValGI 83
43905 GGAAGATGATGG.....AAAGTGGAGGACCCCAAGAAAAA

4398 GAGGGAAAGGAAAAAGGAGGAGGAAAACCGAGGAGGAAGAAG 43947

83 nGlyThrGluValLysArgLysThrThrGluGlnGlnValArgProTyrA 100

43948 AGGTGGGGAGGAAAAGGAGCAGCAGCAGGAGGCAAAATCTCCATCCCCACA 43997

100 rgAra 101

43998 GCAA 44002

seq_name: /SIPS2/acqdata/geneseq/geneseq/NA2001.DAT:AAH45143

seq_documentation_block:

ID
 XX
 AAN45143 SCANDIUM, CDNA; 132 BF.
 XX

XX
XX
XXXXX

DI
XX
00 SEP 2002 (1115C ENERGY)

XX
XX
human protein synthesis, coding sequence

KW X-chromosome-binding mental retardation; lissencephalous disease; ss.

Isolated human brain-expressed X-linked polypeptide used to diagnose PT and treat of dysembryoplasia, hereditary diseases, cancer, tumor, deafness and X-chromosome-binding mental retardation -

P/T deafness and X-chromosome-blinding mental retardation -
 XX
 PS Claim 5; Page 22; 30pp; Chinese.
 XX

CC The present sequence is the coding sequence for a human brain-expressed
 CC X-linked protein (hBex). hBex and its coding sequence are useful in the
 CC diagnosis and treatment of dysembryoplasia, hereditary diseases, cancer,
 CC tumours, deafness, X-chromosome-binding mental retardation and
 CC lissencephalous disease. hBex is also useful for screening mimics,
 CC agonists, or inhibitors, and in peptide fingerprinting identification.
 CC hBex coding sequence can be used as primers or probes, or in producing
 CC gene chips or microarrays.

SQ Sequence 792 BP; 214 A; 172 C; 219 G; 187 T; 0 other;

alignment_scores:

Quality: 94.00 Length: 128
 Ratio: 1.362 Gaps: 2
 Percent Similarity: 53.906 Percent Identity: 28.906

alignment_block:

US-09-327-750D-35 x AAH45143 ..

Align seg 1/1 to: AAH45143 from: 1 to: 792

14 GluLysAspLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerG 30

187 GAACTAGCAGTAACAGTCTCAGCATGGAAATGCCAACCAAGAAATG 236

30 LuGluGluProHisHisLeuGluValGluAsnLys..... 42

237 AAGAAAAG.....GAGCAAGTGTCTAATAAAGGGGCCCTTG 274

43LysProGlyGlyAsnValAr 49

275 GCCTCCCTTGGTGTGTGTAATCTGTGCTAGAGGAATCGTAG 324

49 gArg...LysValArgArgLeuValProAsnPhLeuTrpAlaIleProA 65

325 GCGGTTCCGGTTAGGCAGCCCTCCTGCAGTATAGATGGATATGATGC 374

65 snArgHisValAsp.....ArgAsnGluGlyGlyGluAsp 76

375 ATAGGCTTGGAGAACACACAGGAGGATGAGAGAAGAGAAATATGGAAAG 424

77 ValGlyArgPheValValGlnGlyThrGluValLysArgLysThrThrG1 93

425 ATTGGGAGGGGTGACACAG.....CTGATGGAAGAAGCTGAGGGA 465

93 uGlnGlnValArgProTyrArgArgPheArgThrProGluPro..... 107

466 AAAGCAGTTGAGTCATAGTCTGCGGCGAGTCAGCAGCTGACCCCTCACC 515

108 ..AspAsnHisTyrAspPheCysLeuIlePro 117

516 ATGACCATCATGATGAGTTTGTGCTTATGTC 547

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AA01003

seq_documentation_block:

ID AAC01003 standard; cDNA; 401 BP.

XX AAC01003;

XX 06-OCT-2000 (first entry)

DT Human secreted protein 5' EST, SEQ ID NO: 1001.

DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX P-PSDB; AAG00997.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX Claim 1; SEQ ID 1001; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. An ORF has been identified within the
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 CC derived from 30 different tissues. EST sequences usually correspond
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors.

SQ Sequence 401 BP; 98 A; 98 C; 148 G; 57 T; 0 other;

alignment_scores:

Quality: 92.00 Length: 52
 Ratio: 2.359 Gaps: 2
 Percent Smilarity: 75.000 Percent Identity: 42.308

alignment_block:

US-09-327-750D-35 x AAC01003 ..

Align seg 1/1 to: AAC01003 from: 1 to: 401

1 MetAlaSerLysLysGlnValIleLeu...AspLeuThrValGluLysAs 16

238 ATGGAGTCCAAAGAGGAAGTACTAGCGGCAACAATCTCAACGGGCAAAATGC 287

16 pLysLysAspLysArgGlyGly...LysAlaSerLysGlnSerGluGluG 32

288 CCAACAAGAAACGAAGAGGGAGGAGGCGGCCGCCAGCAATCAAGAG 337

32 LuProHisHisLeuGluValGluAsnLysLysProGlyGlyAsnVal 48

338 AATCCCGCCATTTGGAGGGGTGAAGGCCAGAACCTCGAGGAAATATC 387

49 ArgArg 50

388 AGGCGG 393

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAI20201

seq_documentation_block:

ID AAI20201 standard; DNA; 276 BP.

XX AAI20201;

XX 12-OCT-2001 (first entry)

DE Probe #10134 for gene expression analysis in human cervical cell sample.

THIS PAGE BLANK (USPTO)


```
6024 AGAAACAGGACGACGACGAAGAAGAACACACAAAAAAGAGAGGGA 6073
47 snValArgArgLysValArgLeuValProAsnPheLeuTrpAlaIle 63
6074 AGGAAGAAGGACGAGGAAGA..... 6096
64 ProAsnArgHisValasp.....ArgAsnGluGlyGI 74
|||||.....
6097 CCAAGAAAAACACCAACCCGAAAAAGCAGGAGAGAGGACGAGGAGG 6146
74 yGluAspValGlyArgPheValValGlnGlyThrGluValLys..... 88
|||||.....
6147 AAAAGAAGACCCCAAAAAAAGACGCGACAGACACACACCAAA 6196
89 .....ArgLysThrThrGluGlnGlnValArgProTyrArgArgPhe 102
|||||.....
6197 AACAAACACGAAAAACGACGAACAAACACAAAGAGAGAAAAAGAAA 6246
103 Arg 103
|||
6247 AGA 6249

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-338-907-179
seq_documentation_block:
; Sequence 179, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET-18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 179
; LENGTH: 56520
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2001..2216
; OTHER INFORMATION: exon1
; FEATURE:
; NAME/KEY: exon
; LOCATION: 18196..18265
; OTHER INFORMATION: exon2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 23716..23831
; OTHER INFORMATION: exon3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 25570..25659
; OTHER INFORMATION: exon4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 34668..34758
; OTHER INFORMATION: exon5
; FEATURE:
; NAME/KEY: exon
; LOCATION: 40685..40843
; OTHER INFORMATION: exon6
; FEATURE:
```

```
; NAME/KEY: exon
; LOCATION: 48067..48190
; OTHER INFORMATION: exon7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 50179..54519
; OTHER INFORMATION: exon8
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 5493..5498
; OTHER INFORMATION: AATAAA
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 1991..2008
; OTHER INFORMATION: upstream amplification primer 5-63
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 2505..2525
; OTHER INFORMATION: downstream amplification primer 5-63 , complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 4091..4111
; OTHER INFORMATION: downstream amplification primer 99-622
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 4528..4546
; OTHER INFORMATION: upstream amplification primer 99-622 , complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 5475..5495
; OTHER INFORMATION: downstream amplification primer 99-621
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 5927..5947
; OTHER INFORMATION: upstream amplification primer 99-621 , complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 8127..8144
; OTHER INFORMATION: downstream amplification primer 99-619
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 8560..8578
; OTHER INFORMATION: upstream amplification primer 99-619 , complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 11622..11639
; OTHER INFORMATION: upstream amplification primer 4-76
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 12018..12037
; OTHER INFORMATION: downstream amplification primer 4-76 , complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 11930..11947
; OTHER INFORMATION: upstream amplification primer 4-77
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 12339..12358
; OTHER INFORMATION: downstream amplification primer 4-77 , complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 12915..12932
; OTHER INFORMATION: upstream amplification primer 4-71
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 13317..13334
; OTHER INFORMATION: downstream amplification primer 4-71 , complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 13216..13233
; OTHER INFORMATION: upstream amplification primer 4-72
; FEATURE:
; NAME/KEY: primer_bind
```

```
LOCATION: 13617..13636
FEATURE:
OTHER INFORMATION: downstream amplification primer 4-72, complement
NAME/KEY: primer_bind
LOCATION: 13547..13564
OTHER INFORMATION: upstream amplification primer 4-73
FEATURE:
NAME/KEY: primer_bind
LOCATION: 13962..13981
OTHER INFORMATION: downstream amplification primer 4-73, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 15994..16011
OTHER INFORMATION: downstream amplification primer 99-610
FEATURE:
NAME/KEY: primer_bind
LOCATION: 16463..16480
OTHER INFORMATION: upstream amplification primer 99-610, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 17304..17324
OTHER INFORMATION: downstream amplification primer 99-609
FEATURE:
NAME/KEY: primer_bind
LOCATION: 17814..17832
OTHER INFORMATION: upstream amplification primer 99-609, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 18008..18027
OTHER INFORMATION: upstream amplification primer 4-90
FEATURE:
NAME/KEY: primer_bind
LOCATION: 18423..18442
OTHER INFORMATION: downstream amplification primer 4-90, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 18699..18716
OTHER INFORMATION: downstream amplification primer 99-607
FEATURE:
NAME/KEY: primer_bind
LOCATION: 19164..19182
OTHER INFORMATION: upstream amplification primer 99-607, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 22589..22609
OTHER INFORMATION: downstream amplification primer 99-602
FEATURE:
NAME/KEY: primer_bind
LOCATION: 23111..23129
OTHER INFORMATION: upstream amplification primer 99-602, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 25098..25118
OTHER INFORMATION: downstream amplification primer 99-600
FEATURE:
NAME/KEY: primer_bind
LOCATION: 25657..25674
OTHER INFORMATION: upstream amplification primer 99-600, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 26537..26557
OTHER INFORMATION: downstream amplification primer 99-598
FEATURE:
NAME/KEY: primer_bind
LOCATION: 27022..27040
OTHER INFORMATION: upstream amplification primer 99-598, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 32262..32281
OTHER INFORMATION: downstream amplification primer 99-592
FEATURE:
NAME/KEY: primer_bind
LOCATION: 32823..32841
```

```
OTHER INFORMATION: upstream amplification primer 99-592, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 34215..34233
OTHER INFORMATION: upstream amplification primer 99-217
FEATURE:
alignment_scores:
Quality: 74.50 Length: 82
Ratio: 1.656 Gaps: 2
Percent Similarity: 54.878 Percent Identity: 30.488
alignment_block:
US-09-327-750D-35 x US-09-338-907-179/rev ..
Align seg 1/1 to reverse of: US-09-338-907-179 from: 1 to: 56520
14 GluLysAspLysLysAspLysArgGlyLysAlaSerLysGlnSerG1 30
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
39085 GAAAGCAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 39036
30 uGluGluProHisHisLeuGluValGluAsnLysLysProGlyGlyA 47
: : : : : : : : : : : : : : : : : : : : : : : :
39035 AACGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 38986
47 snValArgArgLysValArgArgLeuValProAsnPheLeuTirAlaIle 63
: : ||| ||| |||||
38985 GAGAGAGAGCGAAGAAAGAAAGG..... 38963
38962 .....AAGAAAGAGGGGAGAGAGACATGAGTTCA 38931
76 pValGlyArgPheValValGlnGlyThrGluValLysArgLysThr 91
: : : : : : : : : : : : : : : : : : : : : : : :
38930 AATTGGGCATACTGAGGCTGAAGGAAACAAACATTTCTAGAAGTAGT 38885
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-996-306-1
seq_documentation_block:
; Sequence 1, Application US/08996306
; Patent No. 5945522
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Chumakov, Ilya
; APPLICANT: Blumefeld, Marta
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,306
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.018A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 1:
```

SEQUENCE CHARACTERISTICS:
LENGTH: 56516 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: GENOMIC DNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Promoter
LOCATION: 1629..1870
IDENTIFICATION METHOD: Proscan
FEATURE:
NAME/KEY: Potential ATG
LOCATION: 1998..2000
FEATURE:
NAME/KEY: Exon 1
LOCATION: 2001..2216
FEATURE:
NAME/KEY: ATG
LOCATION: 2031..2033
FEATURE:
NAME/KEY: TYR phos
LOCATION: 11694..14332
FEATURE:
NAME/KEY: SEQ ID42
LOCATION: 11930..11947
FEATURE:
NAME/KEY: SEQ ID24
LOCATION: 12057..12103
FEATURE:
NAME/KEY: SEQ ID51
LOCATION: compl(12339..12358)
FEATURE:
NAME/KEY: SEQ ID64
LOCATION: 13547..13564
FEATURE:
NAME/KEY: SEQ ID58
LOCATION: 13657..13703
FEATURE:
NAME/KEY: SEQ ID67
LOCATION: compl(13962..13981)
FEATURE:
NAME/KEY: Exon 2
LOCATION: 18196..18265
FEATURE:
NAME/KEY: Exon 3
LOCATION: 23717..23832
FEATURE:
NAME/KEY: Exon 4
LOCATION: 25571..25660
FEATURE:
NAME/KEY: SEQ ID43
LOCATION: 34216..34234
FEATURE:
NAME/KEY: SEQ ID25
LOCATION: 34469..34515
FEATURE:
NAME/KEY: SEQ ID52
LOCATION: compl(34625..34645)
FEATURE:
NAME/KEY: Exon 5
LOCATION: 34669..34759
FEATURE:
NAME/KEY: Exon 6
LOCATION: 40688..40846
FEATURE:
NAME/KEY: Exon 7
LOCATION: 48070..48193
FEATURE:
NAME/KEY: Exon 8
LOCATION: 50182..54523
FEATURE:

NAME/KEY: SEQ ID65
LOCATION: 51149..51168
FEATURE:
NAME/KEY: SEQ ID59
LOCATION: 51448..51494
FEATURE:
NAME/KEY: SEQ ID68
LOCATION: compl(51482..51499)
FEATURE:
NAME/KEY: SEQ ID44
LOCATION: 51596..51613
FEATURE:
NAME/KEY: SEQ ID26
LOCATION: 51612..51658
FEATURE:
NAME/KEY: SEQ ID53
LOCATION: compl(51996..52015)
FEATURE:
NAME/KEY: polyAad signal
LOCATION: 54445..54450
US-08-996-306-1

alignment_scores:
Quality: 71.50 Length: 79
Ratio: 1.663 Gaps: 2
Percent Similarity: 54.430 Percent Identity: 30.380

alignment_block:
US-09-327-750D-35 x US-08-996-306-1/rev ..

Align seg 1/1 to reverse of: US-08-996-306-1 from: 1 to: 56516

17 LysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluPr 33
||||| :||||| |||||||||||||:||||| :||
39077 AAGAAAAG 39028

33 OHIsHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArgA 50
::: :|||:|||||:|||||:|||||:|||||
39027 GGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 38978

50 rGlyValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnArg 66
||| |||||
38977 CGAAGAGAAACAGG..... 38964

67 HisValAspArgAsnGluGlyGlu.....AspValGlyAr 79
:|||||:|||||:|||||:|||||:|||||:|||||
38963AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 38923

79 gPheValValGlnGlyThrGluValLysArgLysThr 91
:|||||:|||||:|||||:|||||:|||||:|||||
38922 TACTGAGGCTGAAGAGACAAACATCTTAGAAGTAGT 38886

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-338-907-1

seq_documentation_block:
; Sequence 1, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CP1CP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22


```

539 .....ACAGATCGTAA 549
71 nGluGlyGlyGluAspValGlyArgPheValValGlnGlyThrGluValL 88
      ::::: |||::: ||| ::::: ||| ::::: ||| ::::: |||
550 AAAAGCAGAAAGAAGAGCTAAACAGAAAGCAGCAGAAAGATAAAGTTA 599
      ::::: |||::: ||| ::::: ||| ::::: ||| ::::: |||
88 ysArgLysThrThrGluGln...GlnValArgProTyrArgArgPheArg 103
      ||::||| ::||| ||| ||| ||| ::::: |||
600 AAGAAAAACAGCTGAACCAACCAACCAACCGCGCTCCAAAAAGCAGAA 649
      ::::: |||::: ||| ::::: ||| ::::: |||
104 ThrProGluPro 107
      . . . . . ||| |||
650 AAACCAAGCTCCA 661

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-258-373-21
seq_documentation_block:
; Sequence 21, Application US/09258373
; Patent No. 6150110
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Xiao, Sheng
; TITLE OF INVENTION: HMG(IY)-LAMAA* FUSION ONCOGENE,
; FILE REFERENCE: B0801/7135/ERP
; CURRENT APPLICATION NUMBER: US/09/258,373
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 60/076,401
; EARLIER FILING DATE: 1998-02-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 1875
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-258-373-21

alignment_scores:
      Quality: 71.00      Length: 118
      Ratio: 1.224      Gaps: 7
Percent Similarity: 49.153      Percent Identity: 29.661

alignment_block:
US-09-327-750D-35 x US-09-258-373-21  ..

Align seg 1/1 to: US-09-258-373-21 from: 1 to: 1875

2 AlaserLysLysGlnValIleLeuAspLeuThrValGluLysAspLysLysL 18
  ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
244 AGCTCGAAGTCAGCCAGCCAGCCCTTGGCCTCCACGAGCAAGAGCGCAC 293
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|

18 asPlyArgGly...GlyLysAlaSerLysGlnSerGluGluProH 34
  ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
294 TCAGAAAGGGGGCCGGGGCAGCGCCGCAAGCAGCCTCCGAAGAGGCCA 343
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|

34 isHisLeuGluGluValGluAsnLysLysPro GlyGlyAsnValArgAr 50
  ||| ||| |||::|::|::|::|::|::|::|::|::|::|::|::|
344 GC.....GAAGTGCACACCTTAAGACAGCTCGGGGCCCAACAGGGA 387
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|

50 gLysValArgArgLeuValProAsnPheLeuTAlaIleProAsnArgH 67
  : ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
388 AGCAAAAAACAAGGTGCTCCCAAGA.....CCCGAANAAC 422
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|

67 isValAspArgAsnGluGlyGlyGluAspValGlyArgPheValValGln 83
  ::::: |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
423 CACCACAACCTCCAGGAAGGAAAC.....CAA 448
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|

84 GlyThrGluValLys.....ArgLysThrThrGluGlnGlnVa 96
  |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
449 GGGCAGACCCCAAAACATGGACAGCAGGAGGAGGCGCATCTCCGAC 498
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|

96 largProTyrArgArgPheArgThrPro.....GluProAspAsnHist 111
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```


Length: 82

```

Ratio: 1.690          Gaps: 5
Percent Similarity: 51.220      Percent Identity: 31.707

alignment_block:
US-09-327-750D-35 x US-09-382-911-5 ..

Align seg 1/1 to: US-09-382-911-5 from: 1 to: 4094

12 ThrValGluLysAspLysLysAspLysArgGlyGly.....Lysal 25
|||||::: ::||| |||||::: |:::
1452 ACTCTGAAGCAACGACGAAGAAGAGAGAGGGCATCTTCNAGAGGAATC 1501
:::|||||::: |:::

25 aserLysGlnSerGluGluGlu..... 32
:|||||::: |:::
1502 CAGCAAGAAGAGCGCTGAGAGGCCGCTGGAGACCCTTCATCATCAGGC 1551
33 .....ProHisHisLeuGluGlu.....ValGluAsnLysLys 43
||| |||::: |:::
1552 CCACCCCTCCCGCTCATGAAGCCCTCTGGTGTGTGAACCCCAAG 1601
44 ProGlyGlyAsnValArgArgLysValArgArgLeuValProAsnPheLe 60
||||| |||::: |:::
1602 AGTGGGGCAACACGGTGCCAAGATC.....ATCCAGTCTTTCT 1642
60 uTrpAlaIleProAsnArgHisVal...AspArgAsnGluGlyGly 74
|||::: |||||::: |:::
1643 CTGGTATCTCAATCCCAGACAAGTCTTCGACCTGAGCCAGGGAGGG 1688

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq;us-08-402-282-3

seq_documentation_block:
; Sequence 3, Application us/08402282
; Patent No. 5476768
; GENERAL INFORMATION:
; APPLICANT: Pearson, Robert E.
; APPLICANT: Dickson, Julie A.
; APPLICANT: Hamilton, Paul T.
; APPLICANT: Little, Michael C.
; APPLICANT: Bever Jr., Wayne F.
; TITLE OF INVENTION: MYCOBACTERIOPHAGE SPECIFIC FOR THE
; TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS COMPLEX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
; ADDRESSEE: Company
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
; ZIP: 07417
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/402,282
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fugit, Donna R.
; REGISTRATION NUMBER: 32,135
; REFERENCE/DOCKET NUMBER: P-3283
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1566 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; NAME/KEY: misc.feature
; LOCATION: 222..425

```

```

; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 451..747
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 747..1109
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1109..2014
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2034..2747
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2747..3109
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3109..3444
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3444..3728
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3731..4855
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4855..5376
; OTHER INFORMATION: /function= "potential coding
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /product= "L5 gp37 homolog"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5382..5747
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5837..6307
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6403..7770
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7770..8006
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8033..8236
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"

```

```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8244..9443
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9450..10244
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10371..10586
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11115..11786
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11917..12741
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12748..14499
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 14771..15154
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15154..15426
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15429..15664
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; US-08-402-282-3
;
alignment_scores:
; Quality: 71.00 Length: 71
; Ratio: 1.775 Caps: 4
; Percent Similarity: 56.338 Percent Identity: 33.803
;
alignment_block:
; US-09-327-750D-35 x US-08-402-282-3
;
Align seg 1/1 to: US-08-402-282-3 from: 1 to: 15664
;
49 ArgArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAs 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13436 CAAATCGGTGGCGCAAGGCGCGCGATGCGCGTGG.....CCGAC 13479
;
65 nArg.His.....ValAspArgAsnGluGly 73
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13480 GCGCTCACCCGACCGGACCTCGAGGTCCGCGTCGATCGGTCCGGGA 13529
;
74 GlyGluAspValGlyArgPheValGlnGlyThrGluValLysArgly 90
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13530 GGAATGAT.....CGAAGCTACCGGCGGACACAGAGAAAAACGGCG 13573
;
90 sThrThrGluGlnValArgProTyrArgArgPheArgThrProGluP 107
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13574 GCGACGACGCGAATCGACGCCGCCGACGACGATCGTGACGTCGATC 13623

```

```
107 roAspAsnHis 110
||
||.....|||
13624 CTGTTGAGCAT 13634

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-508-004-3

seq_documentation_block:
; Sequence 3, Application US/08508004
; Patent No. 5582969
; GENERAL INFORMATION:
; APPLICANT: Pearson, Robert E.
; APPLICANT: Dickson, Julie A.
; APPLICANT: Hamilton, Paul T.
; APPLICANT: Little, Michael C.
; APPLICANT: Beyer Jr., Wayne F.
; TITLE OF INVENTION: MYCOBACTERIOPHAGE SPECIFIC FOR THE
; TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS COMPLEX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
; ADDRESSEE: Company
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
; ZIP: 07417
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/508,004
; FILING DATE: 27-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/402,282
; FILING DATE: 10-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Fugit, Donna R.
; REGISTRATION NUMBER: 32,135
; REFERENCE/DOCKET NUMBER: P-3283
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15664 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 222..425
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 451..747
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 747..1109
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1109..2014
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2034..2747
```

```
;
;
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2747..3109
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3109..3444
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3444..3728
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3731..4855
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4855..5376
; OTHER INFORMATION: /function= "potential coding
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /product= "L5 gp37 homolog"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5382..5747
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5837..6307
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6403..7770
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7770..8006
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8033..8236
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8244..9443
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9450..10244
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10371..10586
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11115..11786
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
```

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11917..12741
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12748..14499
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 14771..15154
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15154..15426
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15429..15664
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
US-08-508-004-3

alignment_scores:
Quality: 71.00      Length: 71
Ratio: 1.775      Gaps: 4
Percent Similarity: 56.338      Percent Identity: 33.803

alignment_block:
US-09-327-750D-35 x US-08-508-004-3  ..
Align seg 1/1 to: US-08-508-004-3 from: 1 to: 15664

49 ArgArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAs 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13436 CGAAATCGGTGGCGGCAAGGCGCGCGCGGTGG.....CCGAC 13479
:|||||:
65 nArg.His.....ValAspArgAsnGluGly 73
|||||:|||||:|||||:|||||:|||||:
13480 GCGCTCACCCGACCGCGACCTCGAGTCGTCGTCGTCGTCGCGGA 13529
74 GlyGluAspValGlyArgPheValGlnGlyThrGluValLysArgly 90
|||||:|||||:|||||:|||||:|||||:
13530 GGGATGAT.....CGAAGCTTACCGCGGCGACACGAAGAAACGGCG 13573
90 sThrThrGluGlnGlnValArgProTyrArgPheArgPheArgThrProGluP 107
:|||||:|||||:|||||:|||||:|||||:
13574 GCGACGACGAGGATCGACGCGCGCGCGACGACGATCGTCAGTCGATC 13623
107 toAspAsnHis 110
||:|||||
13624 CTGTTTCAGCAT 13634

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq.US-08-402-066-3

seq_documentation_block:
; Sequence 3, Application US/08402066
; Patent No. 5612182
; GENERAL INFORMATION:
; APPLICANT: Pearson, Robert E.
; APPLICANT: Dickson, Julie A.
; APPLICANT: Hamilton, Paul T.
; APPLICANT: Little, Michael C.
; APPLICANT: Beyer Jr., Wayne F.
; TITLE OF INVENTION: MYCOBACTERIOPHAGE SPECIFIC FOR THE
; TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS COMPLEX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
```

```
ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
ADDRESSEE: Company
STREET: 1 Becton Drive
CITY: Franklin Lakes
STATE: NJ
COUNTRY: US
ZIP: 07417
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICANT: Becton Dickinson
FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Fugit, Donna R.
REGISTRATION NUMBER: 32,135
REFERENCE/DOCKET NUMBER: P-3283
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 15664 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 222..425
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 451..747
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 747..1109
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1109..2014
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2034..2747
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2747..3109
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3109..3444
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3444..3728
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3731..4855
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: 4855..5376
; OTHER INFORMATION: /function= "potential coding
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /product= "L5 gp37 homolog"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5382..5747
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5837..6307
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6403..7770
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7770..8006
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8033..8236
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8244..9443
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9450..10244
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10371..10586
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11115..11786
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11917..12741
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12748..14499
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 14771..15154
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15154..15426
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15429..15664
```

```
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; US-08-402-066-3

alignment_scores:
    Quality: 71.00      Length: 71
    Ratio: 1.775       Gaps: 4
    Percent Similarity: 56.338   Percent Identity: 33.803

alignment_block:
    US-09-327-750d-35 x US-08-402-066-3  ..

Align seg 1/1 to: US-08-402-066-3 from: 1 to: 15664

49 ArgArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAs 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13436 CGAATCGGTGGCGCAAGGCGCGCGCGATGGCGTGG.....CCGAC 13479

65 nArg.His.....ValAspArgAsnGluGly 73
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13480 GCGCTCACCCCGCAACCGGACCTCGAGTCCGTCGATCGGTCCGGGA 13529

74 GlyGluAspValGlyArgPheValGlnGlyThrGluValLysArgLy 90
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13530 GGGATGAT.....CGAAGCCTACCGCGCGGACACGAGAAACGGCG 13573

90 sThrThrGluGlnValArgProTyrArgPheArgThrProGluP 107
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13574 GCGAGCAGCGGATCGACGCCGACGACGATCGTCACGTCGATC 13623

107 roAspAsnHis 110
|||:|||||
13624 CTGTCAGCAT 13634

seq_name: /cgn2_6/ptodata/2/lna/5a_comb.seq:US-08-402-068-3

seq_documentation_block:
; Sequence 3, Application US/08402068
; Patent No. 5633159
; GENERAL INFORMATION:
; APPLICANT: Pearson, Robert E.
; APPLICANT: Dickson, Julie A.
; APPLICANT: Hamilton, Paul T.
; APPLICANT: Little, Michael C.
; APPLICANT: Beyer Jr., Wayne F.
; TITLE OF INVENTION: MYCOBACTERIOPHAGE SPECIFIC FOR THE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
; ADDRESSEE: Company
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
; ZIP: 07417
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/402,068
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fugit, Donna R.
; REGISTRATION NUMBER: 32,135
; REFERENCE/DOCKET NUMBER: P-3283
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15664 base pairs
```

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 222..425
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 451..747
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 747..1109
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1109..2014
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2034..2747
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2747..3109
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3109..3444
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3444..3728
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3731..4855
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4855..5376
; OTHER INFORMATION: /function= "potential coding
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /product= "L5 gp37 homolog"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5382..5747
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5837..6307
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6403..7770
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7770..8006
```

```

; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8033..8236
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8244..9443
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9450..10244
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10371..10586
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11115..11786
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11917..12741
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12748..14499
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 14771..15154
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15154..15426
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15429..15664
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; US-08-402-068-
alignment_scores:
    Quality: 71.00      Length: 71
    Ratio: 1.775      Gaps: 4
Percent Similarity: 56.338      Percent Identity: 33.803
alignment_block:
US-09-327-750D-35 x US-08-402-068-3      ..
Align seg 1/1 to: US-08-402-068-3 from: 1 to: 15664
49 ArgArgLysValArgLeuValProAsnPheLeuTrpAlaIleProAs 65
||||:|||||
13436 CGAAATCGGTGGCGGCAAGGCGCGGATGCCGTGG.....CGAC 13479
65 nArg.His.....ValaspArgAsnGluGly 73
|||||
13480 GCGCTCACCCCGAACCCGCGACCTCGAGGTCGCGTTCGGGGA 13529
```

74 GlyGluAspValGlyArgPheValValcInGlyThrGluValLysArgLy 90
 |||:||||| ||| :|| ||| |||:|
 13530 GGGAAATGAT.....CGAAGCTACCGGGGACACCAAGAAACGCG 13573
 90 sThrThrGluGlnValArgProTyrArgArgPheArgThrProGluP 107
 : : : : : ||||| |||:|
 13574 GCGACGACGAGGAATCGACGGCCGACGACGACATCGTGACGTCGATC 13623
 107 roAspAsnHis 110
 || :|||
 13624 CTGTTACGAT 13634

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US93-03077-4

```
seq_documentation_block:
; Sequence 4, Application PCT/US9303077
; GENERAL INFORMATION:
; APPLICANT: Board of Regents, The University of Texas System
; APPLICANT: Gaynor, Richard B.
; APPLICANT: Wu, Foon Kin
; TITLE OF INVENTION: PROTEIN CELLULAR FACTOR USEFUL FOR
; TITLE OF INVENTION: REGULATING GENE EXPRESSION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/03077
; FILING DATE: 19930331
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,025
; FILING DATE: April 2, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Kammerer, Patricia A.
; REGISTRATION NUMBER: 29,775
; REFERENCE/DOCKET NUMBER: UTDPD270PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-787-1540
; TELEFAX: 713-749-2679
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; . SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; PCT-US93-03077-4
```

alignment_scores:
 Quality: 70.50 Length: 108
 Ratio: 1.195 Gaps: 6
 Percent Similarity: 54.630 Percent Identity: 28.704

alignment_block:
 US-09-327-750D-35 x PCT-US93-03077-4 ..

Align seg 1/1 to: PCT-US93-03077-4 from: 1 to: 1242

3 SerLysLysGlnVal.....IleLeuAspLeuThrValGluLy 15
 |||||:|||||
 388 TCAAAACAGCAGCTGCACATCTCAATCATCAAGAAATTAAGAGCTAA 437

15 SASpLysLysAspLys.....ArgGlyGlyLysAlaserLysG 28
 |||||:||||| :||| :||| |||:|
 438 AGACAAGGAGAAATGAAATATGTTGCAAAAGCTGAACAAAAAGTTAAAG 487
 28 InSerGluGluProHisHisLeu.....Glu 37
 : : ||||| :|||
 488 AGCTAGAAGAGGAGTTTCAGCATTTTCAACAGGTCCTTGATGGCAAAAGAA 537
 38 GluValGluAsnLysLysProGlyGlyAsnValArgLysValArgAr 54
 |||||:||||| :|||
 538 GAGGTTTGAGAAACACAT.....AGAGAAATATTAAAAA 572
 54 gLeuValProAsnPheLeuTrpAlaIleProAsnArgHisValAspArg 71
 :||| :||| :|||
 573 ACTA.....AATTCATGGTAGAACGCC 595
 71 snGluGlyGlyGluAspValGlyArgPheValValGlnGlyThrGluVal 87
 :||| :|||:|||||:|
 596 AAGAG.....AAAGATCTTGGCCGCTTCAGGTACACATCGATGAACTT 639
 88 LysArgLysThrThrGluGlnGln 95
 :||| :|||
 640 GAAGAAAGAACCGAAGATATTTCAG 663

OM of: US-09-327-750D-35 to: GenEmbl:* out_format : pfs
Date: Mar 11, 2002 3:35 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cpg2_1/USPTO.spool/US09327750/runat_l1032002_101154_20319/app_query.fasta_1.1472
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500
-MINLEN=0 -MAXLEN=2000000000 -USER=US09327750_@CGNJ_1.8673
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-327-750D-35
Query length: 117
Database: GenEmbl:*
Database sequences: 1472140
Database length: -341344837
Search time (sec): 4557.230000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_pr:AK000959	+	355.50	626.28	1.3e-26	1229	AK000959 Homo sapiens cDNA FLJ
gb_pat:AX100231	+	355.50	625.44	1.5e-26	1364	AX100231 Sequence 32 from Pat
gb_pr:HS635019	+	355.50	593.78	8.5e-25	69648	AL0035494 Human DNA sequence
gb_pr:BC0003190	+	175.00	306.91	8.1e-09	793	BC003190 Homo sapiens, p75NTR-a
gb_pr:HUMOGC	+	175.00	305.97	9.1e-09	891	M38188 Human unknown protein fr
gb_pr:AF187064	+	175.00	305.97	9.1e-09	891	AF187064 Homo sapiens p75NTR-as
gb_pr:AF187066	+	154.00	270.34	8.8e-07	700	AF187066 Mus musculus p75NTR-as
gb_ro:AR097440	+	154.00	268.74	1.1e-06	854	AF097440 Mus musculus brain exp
gb_scs:G727208	-	150.00	267.28	1.3e-06	421	G72708 MARC 4953-4954:991939031
gb_ro:AF187065	+	149.50	264.70	1.8e-06	519	AF187065 Rattus norvegicus p75N
gb_pr:HSV351F8	+	138.00	208.08	0.0036	45678	I 270719 Human DNA sequence fr
gb_pr:HS71487	-	124.50	213.50	0.0013	477	G35294 Human sps SHGC-37409, se
gb_pr:HS71487	-	110.50	159.87	1.25	98274	I 295755 Human DNA sequence fr
gb_ro:AR097438	+	107.50	185.74	0.0454	835	AF097438 Mus musculus brain exp
gb_ro:AF051347	+	107.50	185.71	0.0456	838	AF051347 Mus musculus REX-3 mRN
gb_ro:AF097439	+	107.50	177.69	0.1275	2269	AF097437 Mus musculus Bex1 pro
gb_ro:AF097439	+	102.00	176.39	0.1506	785	AF097439 Mus musculus brain exp
gb_pr:AF251053	+	99.50	172.35	0.2528	744	AF251053 Homo sapiens X-linked
gb_pat:AX078272	+	99.50	170.84	0.3070	898	AX078272 Sequence 76 from Paten
gb_pr:HSV870H8	+	99.50	142.25	12.02	31321	I 270233 Human DNA sequence fr
gb_pr:AL133348	-	99.50	140.16	15.70	40584	AL133348 Human DNA sequence f
gb_hg:HSU80B1	+	99.50	140.07	15.88	41029	AL022169 Homo sapiens chromos
gb_pr:HS73599	+	98.50	129.88	58.71	116561	AL0223653 Human DNA sequenc
gb_hg:AC012947	+	98.50	125.19	107.09	208618	AF195953 Homo sapiens membra
gb_hg:AC012947	+	95.00	121.94	162.57	143599	AF012947 Drosophila melanoga
gb_in:AC009183	-	95.00	119.60	219.49	192055	AC009183 Drosophila melanoga
gb_in:AC010124	+	95.00	119.54	221.08	193398	AC010124 Drosophila melanoga
gb_in:AE003685	+	95.00	118.32	258.52	225038	AE003685 Drosophila melanoga
gb_pr:AF183416	+	94.00	163.70	0.7668	642	AF183416 Homo sapiens ovarian g
gb_pr:AF237783	+	94.00	162.02	0.9513	791	AF237783 Homo sapiens brain-exp
gb_pr:AF220189	+	94.00	161.65	0.9972	828	AF220189 Homo sapiens uncharact
gb_pr:HS19814	+	94.00	131.35	48.61	35714	I AC008708 Human DNA sequence f
gb_pr:AC002478	+	91.00	120.55	194.16	70113	AC002478 Human BAC clone GSI -
gb_hg:AC012232	+	91.00	113.22	497.28	174337	I AC012232 Homo sapiens clone
gb_ba:MXAFR2	+	90.50	146.47	6.99	2507	J04157 M.xanthus frzA, frzB an
gb_pr:AP000265	+	90.50	124.52	116.65	38304	AP000265 Homo sapiens genomic
gb_pr:AP000034	+	90.50	116.80	314.19	100000	I AP000034 Homo sapiens genomic
gb_pr:AP000101	+	90.50	116.80	314.19	100000	I AP000101 Homo sapiens genomic
gb_pr:AP000177	+	90.50	116.80	314.19	100000	I AP000177 Homo sapiens genomic
gb_hg:AC026930	+	90.50	114.66	413.39	130443	I AC026930 Homo sapiens chromo

gb_pr:AC005295 + 90.50 114.62 415.44 131070 ! AC005295 Homo sapiens xp2
gb_pr:AP001713 - 90.50 106.95 1.1e+03 340000 ! AP001713 Homo sapiens gen
gb_ba:AE008371 - 90.00 133.00 39.36 11967 ! AE008371 Agrobacterium tum
gb_pr:U72787 + 90.00 129.48 61.75 18510 ! U72787 Homo sapiens cosmid
gb_pr:AL353806 + 90.00 115.06 392.66 111051 ! AL353806 Human DNA sequen

seq_name: gb_pr:AK000959

seq_documentation_block: 1229 bp mRNA PRI 22-FEB-2000
LOCUS AK000959 Homo sapiens cDNA FLJ10097 fis, clone HEMBA1002458, weakly similar
DEFINITION to OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74.

ACCESSION AK000959
VERSION AK000959.1 GI:7021945
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens embryo, 10 weeks whole embryo, mainly head cDNA to
mRNA, clone_lib:HEMBA1 clone:HEMBA1002458.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (sites)
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K.,
Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y. and Oshima,A.

TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished (2000)
REFERENCE 2 (bases 1 to 1229)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2000) to the DDBJ/EMBL/GenBank databases. Takao

Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana,
Kisarazu, Chiba 292-0812, Japan [E-mail:genomics@hri.co.jp,
Tel:81-438-52-3951, Fax:81-438-52-3952]
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.

FEATURES Location/Qualifiers
source 1..1229
organism="Homo sapiens"
db_xref="taxon:9606"
clone="HEMBA1002458"
dev_stage="embryo, 10 weeks"
tissue_type="whole embryo, mainly head"
note="cloning vector: pME18SFL3"

CDS 191..553
note="unnamed protein product"
codon_start=1
protein_id="BAA91443.1"
db_xref="GI:7021946"
translation="MESRELAANNLNGENAGQENEGEQAPTONEERSRLHGCGEGO
KPGGNIRGRVRLRVPNRPWAI PNPHIEHNARDDEVPFGOMELIKRKTRQQMRHY
MRKOTPEPDNDYFCLIP"

BASE COUNT 351 a 243 c 296 g 339 t

ORIGIN
alignment_scores: Quality: 355.50 Length: 120
Ratio: 3.665 Gaps: 3
Percent Similarity: 80.833 Percent Identity: 59.167

alignment_block:
US-09-327-750D-35 x AK000959 ..
Align seq 1/1 to: AK000959 from: 1 to: 1229

details see <http://bacpac.med.buffalo.edu/VECTOR:pcypac2>
IMPORTANT: This sequence is not the entire insert of clone 635G19.
It may be shorter because we only sequence overlapping sections
once, or longer because we arrange for a small overlap between
neighbouring submissions.
The true left end of clone U65A4 (281014) is at 69549 in this
sequence. The true right end of clone U101D3 (285997) is at 100 in
this sequence.

FEATURES

Source	Location/Qualifiers
--------	---------------------

	1..69648
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="X"
	/map="q22.1-22.3"
	/clone="RP4-635G19"
	/clone_lib="RPCI-4"
	439..483
repeat_region	/note="TARI repeat: matches 707..749 of consensus"
repeat_region	501..661
repeat_region	/note="AluB repeat: matches 137..311 of consensus"
repeat_region	662..897
repeat_region	/note="118 copies 2 mer aa 70% conserved"
repeat_region	915..1331
repeat_region	/note="MSTA repeat: matches 1..419 of consensus"
repeat_region	1334..1517
repeat_region	/note="92 copies 2 mer ga 77% conserved"
repeat_region	1927..2020
repeat_region	/note="MIR repeat: matches 109..217 of consensus"
repeat_region	2066..2262
repeat_region	/note="L1M2 repeat: matches 5825..6022 of consensus"
repeat_region	2263..2574
repeat_region	/note="AluSp repeat: matches 1..313 of consensus"
repeat_region	2575..2819
repeat_region	/note="L1M2 repeat: matches 5583..5825 of consensus"
repeat_region	2821..3039
repeat_region	/note="Alu repeat: matches 1..311 of consensus"
repeat_region	3104..3384
repeat_region	/note="L1M2 repeat: matches 5304..5591 of consensus"
repeat_region	3391..4007
repeat_region	/note="L1M1 repeat: matches 5558..6162 of consensus"
repeat_region	4012..4137
repeat_region	/note="Trigger2a repeat: matches 11..123 of consensus"
repeat_region	4141..4252
repeat_region	/note="L1 repeat: matches 2920..3031 of consensus"
repeat_region	4265..4350
repeat_region	/note="HY4 repeat: matches 1..87 of consensus"
repeat_region	4350..4678
repeat_region	/note="MER74A repeat: matches 28..369 of consensus"
repeat_region	4914..5095
repeat_region	/note="MERSA repeat: matches 8..189 of consensus"
repeat_region	5224..5555
repeat_region	/note="L2 repeat: matches 2423..2746 of consensus"
repeat_region	6040..6086
repeat_region	/note="L2 repeat: matches 1676..1720 of consensus"
repeat_region	7413..7478
repeat_region	/note="Alu repeat: matches 239..304 of consensus"
repeat_region	7628..7928
repeat_region	/note="AluB8 repeat: matches 1..316 of consensus"
repeat_region	7952..7987
repeat_region	/note="18 copies 2 mer aa 81% conserved"
repeat_region	7990..8305
repeat_region	/note="AluJo repeat: matches 1..309 of consensus"
repeat_region	8436..8538
repeat_region	/note="MIR repeat: matches 99..212 of consensus"
repeat_region	10836..11442
repeat_region	/note="L1M2 repeat: matches 1357..1972 of consensus"
repeat_region	11443..11571
repeat_region	/note="FLAM.A repeat: matches 2..129 of consensus"
repeat_region	11572..11946
repeat_region	/note="L1M2 repeat: matches 977..1357 of consensus"
repeat_region	11947..12245
repeat_region	/note="AluY repeat: matches 1..301 of consensus"
repeat_region	12246..13098

	/note="L1M2 repeat: matches 275..977 of consensus"
repeat_region	13094..13223
repeat_region	/note="L1M2 repeat: matches -17..113 of consensus"
repeat_region	13523..13820
repeat_region	/note="MLTID repeat: matches 177..505 of consensus"
repeat_region	13823..13932
repeat_region	/note="55 copies 2 mer ct 77% conserved"
repeat_region	13936..13977
repeat_region	/note="MLTID repeat: matches 201..241 of consensus"
repeat_region	13978..14240
repeat_region	/note="AluSg1 repeat: matches 1..270 of consensus"
repeat_region	14245..14534
repeat_region	/note="AluSc repeat: matches 1..290 of consensus"
repeat_region	14535..14721
repeat_region	/note="MLTID repeat: matches 2..201 of consensus"
repeat_region	14724..15604
repeat_region	/note="L1MC/D repeat: matches 4646..5565 of consensus"
repeat_region	15605..15909
repeat_region	/note="AluSc repeat: matches 24..312 of consensus"
repeat_region	16069..16363
repeat_region	/note="AluX repeat: matches 1..291 of consensus"
repeat_region	16482..16641
repeat_region	/note="AluB repeat: matches 136..295 of consensus"
repeat_region	16671..16953
repeat_region	/note="AluX repeat: matches 1..284 of consensus"
repeat_region	17164..17280
repeat_region	/note="L1P repeat: matches 4499..4619 of consensus"
repeat_region	17326..17377
repeat_region	/note="26 copies 2 mer at 79% conserved"
repeat_region	17704..17863
repeat_region	/note="AluSc repeat: matches 118..295 of consensus"
repeat_region	17864..18175
repeat_region	/note="AluB8 repeat: matches 1..318 of consensus"
repeat_region	18176..18298
repeat_region	/note="AluSc repeat: matches 1..118 of consensus"
repeat_region	18848..18873
repeat_region	/note="13 copies 2 mer tg 100% conserved"
repeat_region	18876..19171
repeat_region	/note="AluSg repeat: matches 1..295 of consensus"
repeat_region	19183..19232
repeat_region	/note="25 copies 2 mer ac 100% conserved"
repeat_region	19246..19533
repeat_region	/note="AluY repeat: matches 5..292 of consensus"
repeat_region	19539..19837
repeat_region	/note="MER74B repeat: matches 61..368 of consensus"
repeat_region	20016..20194
repeat_region	/note="MER74A repeat: matches 295..487 of consensus"
repeat_region	20394..20692
repeat_region	/note="AluSg repeat: matches 1..302 of consensus"
repeat_region	20701..20809
repeat_region	/note="LTR1 repeat: matches 682..785 of consensus"
repeat_region	20877..21061
repeat_region	/note="LTR1 repeat: matches 1..195 of consensus"
repeat_region	21186..21636
repeat_region	/note="LTR29 repeat: matches 3..498 of consensus"
repeat_region	21609..21737
repeat_region	/note="MER34 repeat: matches 409..543 of consensus"
repeat_region	21771..22076
repeat_region	/note="AluY repeat: matches 2..308 of consensus"
repeat_region	22878..22974
repeat_region	/note="MIR repeat: matches 92..191 of consensus"
repeat_region	23452..23582
repeat_region	/note="FLAM.C repeat: matches 1..131 of consensus"
repeat_region	23583..23624
repeat_region	/note="21 copies 2 mer ta 76% conserved"
repeat_region	23735..24041
repeat_region	/note="AluSg1 repeat: matches 1..308 of consensus"
repeat_region	24045..24072
repeat_region	/note="14 copies 2 mer tt 89% conserved"
repeat_region	25274..25582
repeat_region	/note="AluB repeat: matches 1..289 of consensus"
repeat_region	25583..25747
misc_feature	/note="match: GSS B63622"


```

|||||.....|
316 CGCGGACTTCCCTTAATTTTCGATGGCCATACCCCAATAGGAGATCAA 365
      :   :   |||.....GlyGluAspValGlyArgPheValValGlnG 84
      :   :   |||.....|
366 TGATGGATGGTGGAGATGGAGATGATATGGAATATTCATGAGAGAGA 415
      :   :   |||.....|
84  lyThrGluVallyArgLysThrThrGluGlnGlnValArgProTyArg 100
      :   :   |||.....|
416 TGAGAGAAATCAGAGAAAATTTAGGAGCTGAGTTCAGCAATTCGCTG 465
      :   :   |||.....|
101 ArgPheArgThrProGluProAspAsnHisTyArgPheValValGlnG 113
      :   :   |||.....|
466 CGTATCCTTATGGGGAGCTCTTAATCACCATGACCATCATGATCAATT 515
      :   :   |||.....|
113 eCysLeuIlePro 117
516 TTGCCTTATGCCT 528

```

seq_name: gb_pr:HUMOGC

seq_documentation_block:

LOCUS HUMOGC 891 bp mRNA 07-MAR-1995
 DEFINITION Human unknown protein from clone PHGR74 mRNA, complete cds.
 ACCESSION M38188 X56942
 VERSION M38188.1 GI:189378
 KEYWORDS
 SOURCE Human ovarian granulosa cell line, cDNA to mRNA.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Rapp,G., Freudenstein,J., Klaudiny,J., Mucha,J., Wempe,F., Zimmer,M. and Scheit,K.H.
 TITLE Characterization of three abundant mRNAs from human ovarian granulosa cells

JOURNAL DNA Cell Biol. 9 (7), 479-485 (1990)
 MEDLINE 91025550
 COMMENT Draft entry and computer-readable sequence for [DNA 9, 479-485 (1990)] kindly submitted by K.H.Scheit, 27-AUG-1990.

FEATURES
 source
 1..891
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="PHGR74"
 /cell_type="granulosa"
 /tissue_type="ovary"
 /note="protein of unknown function"
 /note="protein of unknown function"
 /codon_start=1
 /protein_id="AAA63232.1"
 /db_xref="GI:189379"
 /translation="MANIHQNEEMEQPMQGEEDRPLGGEGHOPAGNRRGQARRLA
 PNFWAIPNRQINDCMGGDDMEIFMEEMREIRKRLQLRNLRLMGLSNHHD
 HHDEFCLMP"
 361..534
 /note="protein of unknown function"
 /codon_start=1
 /protein_id="AAA63233.1"
 /db_xref="GI:189380"
 /translation="MERKALWEEVKATSLQIDGDLADLPLIFDGPPIGRSMGW
 VEMEMIKYSWR"
 251 a 182 c 224 g 234 t

alignment_scores:
 Quality: 175.00 Length: 121
 Ratio: 2.273 Gaps: 4
 Percent Similarity: 63.636 Percent Identity: 36.364

alignment_block:

US-09-327-750D-35 x HUMOGC

Align seg 1/1 to: HUMOGC from: 1 to: 891

```

4  LysLysGlnValIleLeu...AspLeuThrValGluLysAspLysLysAs 19
|||||.....|
297 AAAAAAATCTCATTCATGGCAATATTCACGAGGAAAAACGAAGAGATGGA 346
      :   :   |||.....|
19  pLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluProHisHisL 36
      :   :   |||.....|
347 GCAG.....CCTATGCAGAAATGAGAGGAAGACCGCCCTT 381
      :   :   |||.....|
36  euGluGluValGluAsnLysLysProGlyGlyAsnValArgArgLysVal 52
      :   :   |||.....|
382 TGGGAGGAGGTGAAGCCACCAGCCTCGCAGAAATCGACGGGACAGGCT 431
      :   :   |||.....|
53  ArgArgLeuValProAsnPheLeuTrpAlaIleProAsnArgHisValas 69
|||||.....|
432 CGCGGACTTCCCTTAATTTTCGATGGCCATACCCCAATAGGAGATCAA 481
      :   :   |||.....|
69  pArgAsnGluGly.....GlyGluAspValGlyArgPheValValGlnG 84
      :   :   |||.....|
482 TGATGGATGGTGGAGATGGAGATGATATGGAATATTCATGAGAGAGA 531
      :   :   |||.....|
84  lyThrGluVallyArgLysThrThrGluGlnGlnValArgProTyArg 100
      :   :   |||.....|
532 TGAGAGAAATCAGAGAAAATTTAGGAGCTGAGTTCAGCAATTCGCTG 581
      :   :   |||.....|
101 ArgPheArgThrProGluProAspAsnHisTyArgPheValValGlnG 113
      :   :   |||.....|
582 CGTATCCTTATGGGGAGCTCTCTTAATCACCATGACCATCATGATGAATT 631
      :   :   |||.....|
113 eCysLeuIlePro 117
632 TTGCCTTATGCCT 644

```

seq_name: gb_pr:AF187064

seq_documentation_block:

LOCUS AF187064 891 bp mRNA 11-JUN-2000
 DEFINITION Homo sapiens p75NTR-associated cell death executor (NADE) mRNA, complete cds.

ACCESSION AF187064

VERSION AF187064.1 GI:8452893

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 891)

AUTHORS Mukai,J., Hachiya,T., Shoji-Hoshino,S., Kimura,M.T., Nadano,D., Suvanto,P., Hanaoka,T., Li,Y., Irie,S., Greene,I.A. and Sato,T.A.

TITLE NADE, a p75NTR-associated cell death executor, is involved in signal transduction mediated by the common neurotrophin receptor p75NTR

JOURNAL J. Biol. Chem. 275 (23), 17566-17570 (2000)

MEDLINE 20298829

REFERENCE 2 (bases 1 to 891)

AUTHORS Mukai,J., Hachiya,T., Hoshino,S., Kimura,M., Nadano,D., Suvanto,P., Hanaoka,T., Li,Y., Irie,S. and Sato,T.

TITLE Direct Submission

JOURNAL Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia University, 630 West 168th St., P&S II-451, New York, NY 10032, USA

FEATURES
 Location/Qualifiers
 1..891
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1..891
 /gene="NADE"
 312..647
 /gene="NADE"


```

115 eu1ePro 117
|||||
541 TTATGCCT 548

seq_name: gb_ro:AF097440
seq_documentation_block:
LOCUS AF097440 854 bp mRNA 13-APR-1999
DEFINITION Mus musculus brain expressed x-linked protein 3 (Bex3) mRNA,
complete cds.
ACCESSION AF097440
VERSION AF097440.1 GI:4580593
KEYWORDS
SOURCE
ORGANISM
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 854)
Brown,A.L. and Kay,G.F.
Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
Hum. Mol. Genet. 8 (4), 611-619 (1999)
99172070
Erratum: [[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
2 (bases 1 to 854)
Brown,A.L. and Kay,G.F.
Direct Submission
TITLE
Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
JOURNAL
Location/Qualifiers
FEATURES
source
1..854
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/db_xref="dbEST:AA272375"
/chromosome="X"
/map="near Plp"
/tissue_type="pooled organs"
1..854
/gene="Bex3"
172..546
/gene="Bex3"
/codon_start=1
/product="brain expressed x-linked protein 3"
/protein_id="AAD24431.1"
/db_xref="GI:4580594"
/translation="MANVHOENEEMEOPLONGEDRPVGGCEGHQPAANNNNNNHNN
HHHRRQARLAPFRWAI PNQWMDGLGGDDMEMFMEMREIRRKURELQRLNC
LRILMGLSNHHDDHDFCLMP"
BASE COUNT 237 a 212 c 228 g 177 t
ORIGIN

alignment_scores:
Quality: 154.00 Length: 136
Ratio: 1.925 Gaps: 4
Percent Similarity: 58.824 Percent Identity: 30.147

alignment_block:
US-09-327-750d-35 x AF097440 ..
Align seg 1/1 to: AF097440 from: 1 to: 854
1 MetAlaSerLysLysGlnValIleLeuAspLeuThrValGluLysAspLys 17
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
145 ATAGCAACACAGGAACAAATCTCATCGGCCCATGTCACCAAGGAAA 194
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 sLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluProH 34
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
195 CGAAGAGATG.....GAGCAGCCCTGTCAGATGGACAGAGACC 235
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 iShisLeuGluValGluAsnLysLysProGlyGlyAsn..... 47

```

```

::: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
236 GCCTGTGGAGAGGTGAGGGCCACCAGCTGCTGCAACAAACAAC 285
48 .....ValArgArgLysValArgArg 54
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
286 AACAAACCAACCATAAACACACACACACACCGGAGAGGCCAGGCTCGCG 335
54 gLeuValProAsnPheLeuTrpAlaIleProAsnArgHisValAspArgA 71
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
336 ACTTGCCCTTAATCTCCGATGGCCATTCACCAACAGGCAGATGAATGACG 385
71 snCluGly.....GlyGluAspValGlyArgPheValValGlnGlyThr 85
::: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
386 GGTGGGTGGAGATGAGATGATGAAATGTTTCATGGAGGAGATGAGA 435
86 GluValLysArgLysThrThrGluGlnGlnValValArgProTyrArgArgPh 102
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
436 GAGATCCGGAAGAGCTTAGGAGACTACAGCTGAGAAATTTGTCTACGCAT 485
102 eArgThrProGluProAsnHisTyrAsp.....PheCysL 115
::: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
486 CCTTATGGGGAGCTGCTCTAACCAACCAACGATCACCATGATGAATTTGCC 535
115 eu1ePro 117
|||||
536 TTATGCCT 543

seq_name: gb_sts:G72708
seq_documentation_block:
LOCUS G72708 421 bp DNA STS 08-AUG-2001
DEFINITION MARC 4953-4954:991939031.1 SCF - porcine spleen Sus scrofa STS
genomic, sequence tagged site.
ACCESSION G72708
VERSION G72708.1 GI:15146738
KEYWORDS STS.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
1 (bases 1 to 421)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L. and
Keeler,J.W.
Single nucleotide polymorphism (SNP) discovery in expressed porcine
genes
Unpublished (2001)
JOURNAL
COMMENT
Contact: Freking BA
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4278
Fax: 402 762 4173
Email: freking@email.marc.usda.gov
Primer A: GCAATGGGTGAACACTCTACT
Primer B: CCAGCAGCAATAGACG
STS size: 500
PCR profile:
Hotstart: 95 degrees for 15 minutes
Denature: 95 degrees for 30 seconds
Anneal: 56 degrees
Extension: 68 degrees for 2 minutes
Cycles: 32 to 45
Protocol:
Template: 50-200 ng genomic DNA
Primer: each 20 pmoles
dNTPs: each 88 uM
Taq polymerase: 0.25 units (Qiagen HotStar)
Buffer:
Commercially supplied Qiagen HotStar buffer

```

The STS is derived from PCR amplicons generated from genomic DNA, sequenced from each end using the amplification primers. The

LOCUS HSV351F8 45678 bp DNA PRI 23-NOV-1999
 DEFINITION Human DNA sequence from cosmid V351F8, between markers DXS366 and DXS87 on chromosome X contains ESTs.
 ACCESSION Z70719
 VERSION Z70719.1 GI:1261915
 KEYWORDS X.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 45678)
 AUTHORS Whiteley, M.
 TITLE Direct Submission
 JOURNAL Submitted (09-APR-1995) Sanger Centre, Hinxton, Cambridgeshire, CB10 1RQ, UK. E-mail enquiries: humquery@sanger.ac.uk
 COMMENT IMPORTANT: This sequence is the entire insert of clone V351F8. The true left end of clone V351F8 is at 1 in this sequence. The true right end of clone V351F8 is at 45678.
 V351F8 is from the human chromosome X-specific cosmid library.
 FEATURES
 Location/Qualifiers
 1..45678
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="X"
 /chromosome="X"
 /clone="GHC-351F8"
 /clone_lib="SCCV"
 1..70
 /partial
 /note="Alu repeat: matches 80. .1 of consensus"
 642..930
 /note="Alu repeat: matches 1. .308 of consensus"
 1406..5179
 /note="L1 element fragment"
 5243..5276
 /note="17 copies of 2 mer 82 % conserved"
 5279..5567
 /partial
 /note="Alu repeat: matches 304. .1 of consensus"
 9975..10369
 /note="MSTA element fragment"
 10399..12017
 /note="MSTAR element fragment"
 11318..11978
 /note="THE1BR element fragment"
 11715..11846
 /note="MLTIR element fragment"
 12074..12170
 /note="MSTA element fragment"
 12227..12486
 /note="MSTA element fragment"
 12437..12502
 /note="MSTC element fragment"
 12756..13059
 /partial
 /note="Alu repeat: matches 308. .1 of consensus"
 14478..14650
 /note="THE1B element fragment"
 14491..14536
 /note="MSTA element fragment"
 14715..14831
 /note="THE1B element fragment"
 15099..15533
 /note="L1 element fragment"
 15534..15626
 /note="31 copies of 3 mer 85 % conserved"
 15566..15619
 /note="3 copies of 18 mer 98 % conserved"
 16352..16436
 /note="L1 element fragment"
 16728..16763
 /note="9 copies of 4 mer 94 % conserved"
 16764..17054

/partial
 /note="Alu repeat: matches 308. .1 of consensus"
 19420..19449
 /note="15 copies of 2 mer 87 % conserved"
 complement(19537..20013)
 /note="match: 3' EST N51315 clone 283089"
 20071..20114
 /note="22 copies of 2 mer 98 % conserved"
 20073..20112
 /note="10 copies of 4 mer 100 % conserved"
 20073..20126
 /note="3 copies of 18 mer 87 % conserved"
 23088..23300
 /note="L1 element fragment"
 23311..23490
 /note="MLT1A element fragment"
 23318..23489
 /note="MLT1B element fragment"
 23362..23491
 /note="MLT1C element fragment"
 23507..24407
 /note="L1 element fragment"
 24376..24462
 /note="MSTC element fragment"
 24379..24491
 /note="MSTA element fragment"
 24688..24738
 /note="MSTA element fragment"
 24810..25016
 /note="L1 element fragment"
 25254..25421
 /note="L1 element fragment"
 25439..25696
 /note="MER25 element fragment"
 26284..26343
 /note="MLT1D element fragment"
 26428..26491
 /note="MLT1D element fragment"
 26465..26518
 /note="MLT1E element fragment"
 27446..27667
 /note="L1 element fragment"
 28954..29247
 /partial
 /note="Alu repeat: matches 308. .1 of consensus"
 30160..30325
 /note="2 copies of 83 mer 98 % conserved"
 32588..32875
 /note="match: 5' EST H68599 clone 239077"
 complement(32825..33230)
 /note="match: 3' EST H68239 clone 289077"
 35551..35802
 /partial
 /note="Alu repeat: matches 1. .260 of consensus"
 35804..35839
 /note="18 copies of 2 mer 83 % conserved"
 38625..38705
 /note="MLT1R element fragment"
 38747..38839
 /note="MLT1R element fragment"
 38949..39032
 /note="MLT1E element fragment"
 38965..39076
 /note="MLT1D element fragment"
 BASE COUNT 14260 a 9135 c 9327 g 12956 t
 ORIGIN

alignment_scores:
 Quality: 138.00 Length: 127
 Ratio: 2.000 Gaps: 4
 Percent Similarity: 54.331 Percent Identity: 31.496

alignment_block:

US-09-327-750D-35 x HSV351F8 ..

Align seg 1/1 to: HSV351F8 from: 1 to: 45678

11 LeuThrValGluLysAspLysAspLysArgGlyGlyLysAlaSerLy 27
 19408 TTATCTTTCTAGGAAGAAAAAACAACAGGAAAAAACAACAT 19457
 27 sGlnSerGluGluProHisHisLeuGluGlu.....ValGluA 41
 19458 GGAAATGTCCCAAGGAAAAACAAGTTGTGGAGAGGCCCAAGTCGAGA 19507
 41 sn.....LysLysProGlyGlyAsn 47
 19508 ATNAGCCCCCGCTTTAGGAGGTGCTGAATACAGGAGCTGGAGGAAT 19557
 48 ValArgArgLysValArgArgLeuValProAsnPheLeuTrpAlaIlePr 64
 19558 GTTAAAGGGGTTTGGGCTCCACCTGCCCGGGTTTGGAGAGGATGTGCC 19607
 64 oAsnArgHisValAspArg.....AsnGluGlyGlyGluAspV 77
 19608 CAATAGGCTTGTGCATACATTCATATGATAGATGGAGATGGAGATGATA 19657
 77 alGlyArgPheValValGlnGlyThrGluValLysArgLysThrGlu 93
 19658 TGGACGGTTCATGGAGGAGATGAGAGAGCTAAGGAGGAATATTAGGAA 19707
 94 GlnGlnValArgProTyrArgArgPheArgThrProGluPro..... 107
 19708 CTTCAAGTTAGGTAGCTGCGCATCTTATAGGCGACCTCTCTCACCA 19757
 108 AspAsnHisTyrAspPheCysLeuIlePro 117
 19758 TGATCATCATGATGATGAGTTTGCCTTATGCTTATGCT 19788

seq_name: gb_sts:G35294

seq_documentation_block: 477 bp DNA 02-OCT-1997
 LOCUS human STS SHGC-37409, sequence tagged site.
 DEFINITION G35294
 ACCESSION G35294
 VERSION G35294.1 GI:2459462
 KEYWORDS STS; STS sequence; primer; sequence tagged site.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 MYERS R.M.
 JOURNAL Unpublished (1997)

Contact: Richard M. Myers
 Stanford Human Genome Center (SHGC)
 Stanford University School of Medicine
 Department of Genetics, M-344, Stanford, CA 94305, USA
 Tel: 4157259687
 Fax: 4157259689
 Email: myers@shgc.stanford.edu

Primer A: AACATCTTTCATGAAAGTTGATG

Primer B: CTTTGGCATCTTCTGCAA

STS size: 106

PCR Profile:

Initial incubation: 95 degrees C for 10 minutes

Denaturation: 94 degrees C for 15 seconds

Annealing: 60 degrees C for 30 seconds

Polymerization: 72 degrees C for 23 seconds

PCR Cycles: 30

Thermal Cycler: Perkin Elmer 9600

Protocol:

Template: 25 ng each 1 um
 Primer: each 200 um
 dNTPs: each 200 um
 Amplifaq Gold Polymerase: 0.07 units/ul
 Total Vol: 5 ul

Buffer: MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from N51315
 -- Washington University/Merck EST sequence.

FEATURES

source
 1..477
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 STS
 44..149
 primer_bind
 44..67
 primer_bind
 complement(130..149)
 BASE COUNT 153 a 115 c 79 g 130 t
 ORIGIN

alignment_scores:

Quality: 120.50 Length: 83
 Ratio: 2.410 Gaps: 2
 Percent Similarity: 60.241 Percent Identity: 36.145

alignment_block:

US-09-327-750D-35 x G35294/rev ..

Align seg 1/1 to reverse of: G35294 from: 1 to: 477

42 LysLysProGlyGlyAsnValArgArgLysValArgArgLeuValProAs 58
 474 CAGGAGCTGGAGGAATGTTAAAGGGTTTGGCTCCACCTGCCCGGG 425
 58 nPheLeuTrpAlaIleProAsnArgHisValAspArg.....A 71
 424 TTTTGGAGAGCATGTCCTCAATAGGCTTGTGATAACATTTGATATGATAG 375
 71 snGluGlyGlyGluAspValGlyArgPheValValGlnGlyThrGluVal 87
 374 ATGAGATGAGATGATATGGAACGGTTTCATGGAGGATGAGAGAGCTA 325
 88 LysArgLysThrThrGluGlnGlnValArgProTyrArgArgPheArgTh 104
 324 AGGAGGAATTTAGGGAACCTTCAGTTGAGGTACAGTCTGGGCATTTCTAT 275
 104 rProGluPro.....AspAsnHisTyrAspPheCysLeuIlePro 117
 274 AGGGACCTTCCTCACCATGATCATGATGATGATGATGATGATGATGCT 226

seq_name: gb_pr:HS714B7

seq_documentation_block:

LOCUS HS714B7 98274 bp DNA PRI 12-DEC-1999
 DEFINITION Human DNA sequence from clone CTA-714B7 on chromosome 22q12.2-13.2
 Contains pseudogene similar to part of COX7B (cytochrome c oxidase
 subunit VIIb), a novel pseudogene, ESTs, STS and GSSs, complete
 sequence.

ACCESSION Z99755
 VERSION Z99755.1 GI:3036782
 KEYWORDS HTG; COX7B; cytochrome c oxidase.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 98274)
 AUTHORS Connor, R.
 TITLE Direct Submission

/note="MIR repeat: matches 102. .143 of consensus"
25378. .25555
/note="L2 repeat: matches 2330. .2519 of consensus"
25673. .25910
/note="MIR repeat: matches 4. .255 of consensus"
26722. .26798
/note="L2 repeat: matches 2572. .2649 of consensus"
27040. .27107
/note="L2 repeat: matches 2489. .2500 of consensus"
27108. .27374
/note="AluY repeat: matches 41. .304 of consensus"
27375. .27471
/note="L2 repeat: matches 2372. .2489 of consensus"
27651. .27945
/note="AluX repeat: matches 1. .299 of consensus"
28235. .28876
/note="L2 repeat: matches 1113. .1817 of consensus"
28975. .29097
/note="MER96 repeat: matches 42. .173 of consensus"
29173. .29196
/note="L2 copies 2 mer aa 100 conserved"
29832. .30260
/note="match: GSS: Em:AQ075670"
30392. .30597
/note="MIR repeat: matches 35. .252 of consensus"
30845. .31010
/note="MER91B repeat: matches 2. .162 of consensus"
31056. .31094
/note="MIR repeat: matches 218. .256 of consensus"
31728. .31791
/note="L6 copies 4 mer tgtg 76 conserved"
31858. .31952
/note="MIR repeat: matches 48. .153 of consensus"
complement(33437. .33589)
/note="match: STS: Em:H5731"
33543. .33633
/note="MIR repeat: matches 84. .165 of consensus"
33785. .33966
/note="MER91A repeat: matches 1. .185 of consensus"

alignment_scores:
Quality: 114.50 Length: 66
Ratio: 2.602 Gaps: 3
Percent Similarity: 66.667 Percent Identity: 39.394

alignment_block:

US-09-327-750D-35 x HS714B7

Align seg 1/1 to: HS714B7 from: 1 to: 98274

55 LeuValProAsnPheLeuTTPAlaIleProAsnArgHisValAsp..... 69
||||:|||||
62887 TTGGCCCTTAATTTTCATGGCCATACCAATAGGAGGTCATGATGG 62936
70 ArgAsnGluGlyGlyGluAlaArgPheValValcInGlyThrG 86
||||:|||||
62937 GATCAATAGGTGGAGATGATATGGAATGTTTCATGGAGGATGAG 62986
86 luValLysArgLys.....ThrThrGluGlnGlnValArgPro 98
||||:|||||
62987 GAATCAGGAGAGATAAGGAGCTACAAATTGAGGAATGCTGTGTAFCCT 63036
99 TyrArgArgPheArgThrProGluProAspAsnHisTyrAspPheCys 114
||||:|||||
63037 TATGGGAAGCTGTGTAATCCCAT...GACCATCATGATGAATTTTC 63081

seq_name: gb_ro:AF097438

seq documentation_block:
LOCUS AF097438 835 bp mRNA ROD 11-APR-1999
DEFINITION Mus musculus brain expressed x-linked protein 1 (Bex1) mRNA,
complete cds.
ACCESSION AF097438

AF097438.1 GI:4580589
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 835)
Brown, A. L. and Kay, G. F.
Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
Hum. Mol. Genet. 8 (4), 611-619 (1999)
99172070
Erratum: [[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
2 (bases 1 to 835)
Brown, A. L. and Kay, G. F.
Direct Submission
Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
Location/Qualifiers
1. .835
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="X"
/map="near plp"
/tissue_type="blastocyst"
1. .835
/gene="Bex1"
205. .591
/gene="Bex1"
/codon_start=1
/product="brain expressed X-linked protein 1"
/protein_id="AAD24429.1"
/db_xref="GI:4580590"
/translation="MESKDDQVKNLNMENDHOKBEKEKPDQTIIRREPAVALISEAG
KNCAPRGRRFRVROPJAHYRWDLQMRVGPQGRMREENVQFGDVRQLMEKLRR
QLSHSLRAVSTDPHPHHDHDEFLP"
BASE COUNT 238 a 184 c 246 g 167 t
ORIGIN

alignment_scores:
Quality: 107.50 Length: 129
Ratio: 1.473 Gaps: 6
Percent Similarity: 56.589 Percent Identity: 30.233

alignment_block:
US-09-327-750D-35 x AF097438

Align seg 1/1 to: AF097438 from: 1 to: 835

1 MetAlaSerLysLysGlnValIleLeuAspLeuThrValGluLysAsp.. 16
||||:|||||
205 ATGGAGTCCAAAGATCAAGCGCTGAAATAATCTCAACATGGAGATGACCA 254
17LysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluG 32
||||:|||||
255 TCACAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 304
32 luProHisHisLeu.....GluGluValGluAsnLysLysPro... 44
||||:|||||
305 AGCCAGCTGTGGCCCTGATCTCCGAGGCTGGCAAAACTGTGCGCCTAGA 354
45 GlyClyAsnValArgArgLysValArgLysValProAsnPheLeuTr 61
||||:|||||
355 GGAGGTGCGAGGCGGTTCGCGGTTCGCGCCTGCTCACTATAGATG 404
61 pAlaIleProAsnArgHisValAspArgAsnGluGly.....GlyG 75
||||:|||||
405 GGACCTGATCGAGG...GTGGGGAGCCCCAGGGAAGGATGAGAGAGG 451
75 luAspValGlyArgPheValValcInGlyThrGluValLysArgLysThr 91
||||:|||||

THIS PAGE BLANK (USPTO)

452 AGAACGTACAGAGGTTTGGGGTGATGTGACACAGCTCATGGAGAAGCTG 501
 92 ThrGluGlnGlnValArgProTyrArgPheArgThrProGluPro... 107
 502 AGGAAAGGAGCTGACCCAGCTCGGGCGTTAGCACTGACCCGCC 551

108AspAsnHisTyrAspPheCysLeuIlePro 117
 552 TCATCATGACCACCATGATGAGTTTTCCTCATGCC 588

seq_name: gb_ro:AF051347

seq_documentation_block:
 LOCUS AF051347 838 bp mRNA ROD 01-OCT-1998
 DEFINITION Mus musculus REX-3 mRNA, complete cds.
 ACCESSION AF051347
 VERSION AF051347.1 GI:3510642
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Faria,T.N., LaRosa,G.J., Wilen,E., Liao,L. and Gudas,L.J.
 TITLE Characterization of genes which exhibit reduced expression during the retinoic acid-induced differentiation of F9 teratocarcinoma cells: involvement of cyclin D3 in RA-mediated growth arrest
 JOURNAL Mol. Cell. Endocrinol. 143 (1-2), 155-166 (1998)
 MEDLINE 99021197

REFERENCE
 AUTHORS Faria,T.N., LaRosa,G., Wilen,E., Liao,L. and Gudas,L.J.
 TITLE Direct Submission
 JOURNAL Submitted (26-FEB-1998) Pharmacology, Cornell University Medical College, 1300 York Avenue, New York, NY 10021, USA

FEATURES
 source

1..838
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /cell_line="F9 teratocarcinoma"
 229..579
 /note="retinoic acid reduced gene expression"
 /codon_start=1
 /product="REX-3"
 /protein_id="AAC61929.1"
 /db_xref="GI:3510643"
 /translation="MENDHQKKEEKEKPQDTIRREPAVALISEAGKNCAPRGRRRF
 RYRPIAHYRDLMQRVGEPQGRMEENVRFPGDVQIMKELRERQLSHSLRAVSTD
 PPHDHDHDFCLMP"
 BASE COUNT 238 a 181 c 251 g 168 t
 ORIGIN

alignment_scores:

Quality: 107.50 Length: 129
 Ratio: 1.473 Gaps: 6
 Percent Similarity: 56.589 Percent Identity: 30.233

alignment_block:

US-09-327-750D-35 x AF051347

Align seg 1/1 to: AF051347 from: 1 to: 838

1 MetAlaSerLysLysGlnValIleLeuAspLeuThrValGluLysAsp.. 16
 193 ATGAGTCCAAAGATCAATCGCGTGAATAATCTCAACATGGAGATGACCA 242
 17LysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluG 32
 243 TCAGAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 292
 32 luProHisHisLeu.....GluGluValGluAsnLysLysPro... 44
 293 AGCCAGCTGTGGCCCTGATCTCCGAGGCTGGCAAAACTCTGCGCCTAGA 342

45 GlyGlyAsnValArgArgLysValArgLeuValProAsnPheLeuTr 61
 343 GGAGTCCAGCGCGTTCCGGGTCGCGACGCCCATCGCTCACTATAGATG 392
 61 PaLaIleProAsnArgHisValAspArgAsnGluGly.....GlyG 75
 393 GGACCTGATGCAGAGG...GTTGGGAGCCCCAGGGAAGGATGAGAGAGG 439
 75 LuAspValGlyArgPheValValGlnGlyThrGluValLysArgLysThr 91
 440 AGAACGTACAGAGGTTTGGGGGTGATGTGACACAGCTCATGGAGAAGCTG 489
 92 ThrGluGlnGlnValArgProTyrArgArgPheArgThrProGluPro... 107
 490 AGGAAAGGAGGAGCTGAGCCACAGCTCGGGCGTTAGCACTGACCCGCC 539
 108AspAsnHisTyrAspPheCysLeuIlePro 117
 540 TCATCATGACCACCATGATGAGTTTTCCTCATGCC 576

```

33 roHishLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 49
|||||
246 CCCACCATCTGGAAGAGTTGAAACAAGAGCTGGGGAAATGTCCGA 295
|||||
50 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 66
|||||
296 AGGAAAGTCAGGCGACTTGCTGCTACTTCTGTGGCCATACCAATAG 345
|||||
66 ghisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValValG 83
|||||
346 GCATGTTGATCGCAATGAGGGGGAGAGCATGTTGGAGATTGCTGTC 395
|||||
83 InGlyThrGluValLysArgLysThr 91
|||||
396 AGGGAACAGAGTCAAGAGAAAGACT 421
|||||

```

seq_name: gb_est1:AW251884

```

seq_documentation_block:
LOCUS      AW251884          540 bp      mRNA          EST          17-DEC-1999
DEFINITION UI-R-BJ0-adn-b-06-0-UI.s1 UI-R-BJ0 Rattus norvegicus cDNA clone
            UI-R-BJ0-adn-b-06-0-UI 3', mRNA sequence.
ACCESSION  AW251884
VERSION    AW251884.1 GI:5595475
KEYWORDS   EST.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 540)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
JOURNAL   discovery
MEDLINE   Genome Res. 6 (9), 791-806 (1996)
COMMENT   9704477
            Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. The sequence tag present in the cDNA between the NotI site
            and the oligo-dT track served to identify it as a clone from the
            normalized atrium at 16.5 dpc library cDNA Library Preparation:
            M.B. Soares Lab Clone distribution: clones will be available
            through Research Genetics (www.resgen.com)
            Seq primer: M13 Forward
            POLYA=yes.

```

FEATURES

```

Location/Qualifiers
1..540
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BJ0-adn-b-06-0-UI"
/clone_lib="UI-R-BJ0"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pMT3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-BJ0
library is a subtracted library derived from the UI-R-AAL,
UI-R-AB1, UI-R-AC1, UI-R-AD1, UI-R-AE1, UI-R-AF1, and
UI-R-AG1 libraries. These libraries represent tissues from
rat atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal
at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV
canal at 15 dpc, and ventricle at 13 dpc. The tag is a
string of 5-6 nucleotides present between the Not I site
and the oligo-dT track. The library was constructed as
described by Bonaldo, Lennon and Soares, Genome Research

```

```

6: 791-806, 1996.
TAG_LIB=UI-R-BJ0
TAG_TISSUE=atrium at 16.5 dpc
TAG_SEQ=GATTC"
BASE COUNT      146 a      120 c      94 g      180 t
ORIGIN
alignment_scores:
Quality: 459.00      Length: 97
Ratio: 5.044      Gaps: 0
Percent Similarity: 93.814      Percent Identity: 86.598
alignment_block:
US-09-327-750D-35 x AW251884/rev ..
Align seg 1/1 to reverse of: AW251884 from: 1 to: 540
21 ArgGlyGlyLysAlaSerLysGlnSerGluGluGluProHisHisLeuCl 37
:::|||||
538 AAAGGTGGGAGGCGCTCCAAACAAGTGAAGAGAATCCACCATCTGGA 489
|||||
37 uGluValGluAsnLysLysProGlyGlyAsnValArgArgLysValArgA 54
|||||
488 ACAGGTTGAAACAAGACCGCTGGGGCAATGTGAGGAGGAAGTCAGGC 439
|||||
54 rGluValProAsnPheLeuTrpAlaIleProAsnArgHisValAspArg 70
|||||
438 GACTTGTGCTTAACCTTCTATGGGCCATACCTAANTAGGCATGTTGATC 389
|||||
71 AsnGluGlyGlyGluAspValGlyArgPheValGlnGlyThrGluVa 87
:::|||||
388 ACTGAAGGGGGAGAGAGGCTTGGGAGATTCTGAGGACGGTGATGGAAGC 339
|||||
87 lLysArgLysThrThrGluGlnGlnValArgProTyrArgArgPheArgT 104
:::|||||
338 CAAGAGAAAGTCTAAGGAGCAACAGATGAGGCTTACACGCGTTTCCGAA 289
|||||
104 hrProGluProAspAsnHisTyrAspPheCysLeuIlePro 117
|||||
288 CCCCTGAACCTGACAATCATTAGGACTTTTGCTCATACCT 248
|||||

```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 465)

REFERENCE
AUTHORS
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, I., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE
JOURNAL
COMMENT
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouse@wustl.edu
This clone is available royalty-free through INL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:229324

Seq primer: ETPRimer
High quality sequence stop: 328.
Location/Qualifiers
1. 465
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:367892"
/clone_lib="Soares mouse embryo NBMEl3.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/note="Vector: pMT3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGCGCGCGGAAATTTTTTTTTTTTTTTT 3'], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne State Univ., from 2]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 147 a 90 c 156 g 72 t
ORIGIN

alignment_scores:
Quality: 469.50 Length: 99
Ratio: 4.840 Gaps: 1
Percent Similarity: 97.980 Percent Identity: 96.970

alignment_block:
US-09-327-750d-35 x W54487 ..

Align seg 1/1 to: W54487 from: 1 to: 465

1 MetAlaSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
|||||
168 ATGGCATCCAAATTTAAACAAGTCATACCTGGATCTCCTGAGGAGAA 217

16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGlu 33
|||||
218 CAAAAAGACAAAGAGGTGGGAGGCCCTCCAAACAAGTGAAGAAGAAC 267

33 roHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 49
|||||
268 CCCACCATCTGGAAGAGGTGAAACAAGAGCCTGGGGGAATGTCCGA 317

50 ArgLysValArgArgLeuValProAsnPheLeuTirPalalIleProAsn 66
|||||

318 AGAAGTC. AGCGCACTTGTGCTAACTTTCTCTGGGCCATACCAATAG 366

66 qHisValAspArgAsnGlyGlyValGluAspValGlyArgPheValValC 83
|||||

367 GCATGTTGATCGCAATGAGGGGAGAGGATGTTGGGAGATTGTAGTGC 416
|||||

83 InGlyThrGluValLysArgLysThrThrGluGlnGlnValArgPro 98
|||||

417 AGGGAACAAGTCACAGAAAGACTACGAGCAGCAGGTGAGCCT 463

seq_name: gb_estl:AL118340

seq_documentation_block:
LOCUS AL118340 422 bp mRNA EST 23-SEP-1999
DEFINITION V9112b26 Beddington mouse dissected endoderm Mus musculus CDNA
clone 528_13N20 5', mRNA sequence.

ACCESSION AL118340
VERSION AL118340.1 GI:5920179
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 422)
REFERENCE
AUTHORS
Harrison, S.M., Dunwoodie, S.L., Arkell, R.M., Lehrach, H. and Beddington, R.S.
TITLE
Isolation of novel tissue-specific genes from cDNA libraries representing the individual tissue constituents of the gastrulating mouse embryo
Development 121 (8), 2479-2489 (1995)
MEDLINE 95401865
COMMENT
Contact: Ruiz, P., Lehrach, H. and Avner, P.
EEC Mouse Transcript Mapping Consortium
Genoscope - CNS
2, rue Gaston Cremieux, CP 5706, 91057 Evry CEDEX, France
Email: pavner@pasteur.fr
clone available from Ressourcenzentrum, Heubnerweg 6, D-14059 Berlin, Germany. Web site http://www.rzpd.de
Seq primer: CCGTCGCGAATTCGCGGT.
Location/Qualifiers
1. 422
/organism="Mus musculus"
/strain="C57B16 x DBA"
/db_xref="taxon:10090"
/clone="528_13N20"
/clone_lib="Beddington mouse dissected endoderm"
/tissue_type="dissected endoderm"
/dev_stage="7.5 dpc"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. ~ High quality sequence only submitted. ~ Average insert size: 1.2 kb (range: 0.2 - 2.kb)"

BASE COUNT 128 a 86 c 143 g 65 t
ORIGIN

alignment_scores:
Quality: 463.50 Length: 92
Ratio: 5.093 Gaps: 1
Percent Similarity: 98.913 Percent Identity: 98.913

alignment_block:
US-09-327-750d-35 x AL118340 ..

Align seg 1/1 to: AL118340 from: 1 to: 422

1 MetAlaSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
|||||

146 ATGGCATCCAAATTTAAACAAGTCATACCTGGATCTCCTGAGGAGAA 195

16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGlu 33
|||||

196 CAAAAAGACAAAGAGGTGGGAGGCCCTCCAAACAAGTGAAGAAGAAC 245


```

100 ArgArgPheArgThrProGluProAspAsnHisTyrAspPheCysLeu1116
302 ACCGTTTCGAACCCCTGAACCTGACATCATACGACTTTTGCCTCAT 253
116 ePro 117
1111
252 ACCT 249

```

seq_name: gb_est1:AAL17429

```

seq_documentation_block:
LOCUS AAL17429 312 bp mRNA EST 15-NOV-1996
DEFINITION mn2le12.r1 Beddington mouse embryonic region Mus musculus cDNA
clone IMAGE:538606 5' similar to SW:HG74_HUMAN Q00994 OVARIAN
GRANULOSA CELL 13.0 KD PROTEIN HGR74. ;, mRNA sequence.
ACCESSION AAL17429
VERSION AAL17429.1 GI:1672442
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 312)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:325542
Seq primer: -40m13 ET
High quality sequence stop: 287.
Location/Qualifiers
1..312
/organism="Mus musculus"
/strain="C57BL/6 x DBA"
/db_xref="taxon:10090"
/clone="IMAGE:538606"
/clone_lib="Beddington mouse embryonic region"
/sex="pooled"
/tissue_type="embryo"
/dev_stage="7.5dpc"
/lab_host="DH12S"
/Note="Organ: whole embryo; Vector: pCMV-Sport; Site_1:
Sali; Site_2: NotI; Cloned unidirectionally. Primer:
Oligo dt. Gastrulating embryos were collected at 7.5dpc
from C57BL6 x DBA matings, excluding embryos that had
developed head folds and all extraembryonic tissues.
Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).
Referenced in Development 121, 2479-2489 (1995)."
BASE COUNT 110 a 56 c 92 g 54 t
ORIGIN

```

```

alignment_scores:
Quality: 503.50 Length: 100
Ratio: 5.086 Gaps: 1
Percent Similarity: 99.000 Percent Identity: 98.000

```

alignment_block:

US-09-327-750D-35 x AAL17429 ..

Align seg 1/1 to: AAL17429 from: 1 to: 312

```

1 MetAlSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
|||||
12 ATGCATCCAAATTTAAACAAGTCATCTACTGGATCTCACTGTGGAGAAGA 61
|||||
16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGlu 33
|||||
62 CAAAAAGACAAAAAGGTGGAGGCTCCAAACAAGTGAAGAAGAAC 111
|||||
33 roHisHisLeuGluValGluAsnLysLysProGlyGlyAsnValArg 49
|||||
112 CCCACCATCTCGGAAGAGTTGAAAACAAGAGCCTGGGGGAATGTCCGA 161
|||||
50 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 66
|||||
162 AGCAAAGTCAGCGACTTGTGCCTAACTTCTCTGGGCCCATACCTAATAG 211
|||||
66 gHisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValValG 83
|||||
212 GCATGTTGATCGCAATGAAGGGGAGAGGATGTTGGGAGATTTGTAGTGC 261
|||||
83 InGlyThrGluValLysArgLysThrThrGluGlnGlnValArgProTyr 99
|||||
262 AGGAACAGAAAGTCAGAGAAAGACTACGAGCAGCAGCGTGGAGCCCTAC 311
|||||
seq_name: gb_est2:W35893

```

seq_documentation_block:

```

LOCUS W35893 472 bp mRNA EST 14-MAY-1996
DEFINITION mc53g09.r1 Soares mouse embryo NBME13.5 14.5 Mus musculus cDNA
clone IMAGE:352288 5' similar to SW:HG74_HUMAN Q00994 OVARIAN
GRANULOSA CELL 13.0 KD PROTEIN HGR74. [1] ;, mRNA sequence.
ACCESSION W35893
VERSION W35893.1 GI:1317733
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 472)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:224088
Seq primer: EPrimer
High quality sequence stop: 441.
Location/Qualifiers
1..472
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:352288"
/clone_lib="Soares mouse embryo NBME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/Note="Vector: pTT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand
CDNA was primed with a Not I - oligo(dT) primer [5,

```

FEATURES
source

```

|||||
428 CCCACCATCTCGAAGAGGTTGAAACAAGAGCCTGGGGCAATGTGTCAGG 379
|||||
50 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 66
|||||
378 AGGAAAGTCAGGCGACTGTGTCCTAACTTTCTATGGGCCATACCTAATAG 329
|||||
66 gHisValAspArgAsnGluGlyGluAspValGlyArgPheValValG 83
|||||
328 GCATGTTGATCATCAGTGAAGGGGAGGAGGTGGAGAGATTGCTAGGGC 279
|||||
278 AGGTGATGGAAGCCAGAGAAAGTCTAAGGAGCAACAGATGAGGCGCTTAC 229
|||||
83 InGlyThrGluValLysArgLysThrThrGluGlnValArgProTyr 99
|||||
100 ArgArgPheArgThrProGluProAspAsnHisTyrAspPheCysLeuII 116
|||||
228 ACCGCTTTCGAACCCCTGAACCTGAACATCAATAGGACTTTTGCCTCAT 179
|||||
116 ePro 117
|||||
178 ACCT 175

```

seq_name: gb_est2:BI289546

```

seq_documentation_block: 643 bp mRNA EST 19-JUL-2001
LOCUS BI289546
DEFINITION UI-R-DK0-cff-c-12-0-UI-s1 UI-R-DK0 Rattus norvegicus cDNA clone
UI-R-DK0-cff-c-12-0-UI 3', mRNA sequence.
ACCESSION BI289546
VERSION BI289546.1 GI:14947228
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 643)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized rat kidney pool library cDNA Library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com) the following repetitive
elements were found in this cDNA sequence: 1-51,
>AT_rich#Low_complexity
Seq primer: M13 Forward
POLYA=Yes.

```

FEATURES
source

```

Location/Qualifiers
1..643
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DK0-cff-c-12-0-UI"
/clone_lib="UI-R-DK0"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/note="vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-DK0

```

library is a subtracted library derived from a mixture of five individually tagged normalized rat libraries: brain-nRBP (20%), heart-nRHP (20%), kidney-nRKP (20%), aorta-nRAP (20%), and placenta-nRPP (20%). Each original library was constructed from a mixture of equal amounts of RNA from seven different developmental time-points: embryonic day 17, embryonic day 19, embryonic day 21, adult day 1, adult day 12, adult day 75, and adult day 200. (Exception: the aorta pool does not contain embryonic day 17 RNA and the placenta pool contains only the three embryonic stages). Each library was normalized individually according to the procedure described by Bonaldo, Lennon & Soares (Genome Research Genome 6: 791-806, 1996). For construction of the DK0 subtracted library, plasmid DNA from each of the five individually tagged normalized libraries was mixed in the proportions specified above and electroporated into competent bacteria for production of single-stranded circular DNA representing the pool of libraries. Single-stranded circular DNA representing these five normalized libraries was then used as a tracer in a subtractive hybridization with a driver (PCR amplified inserts from a plasmid DNA template preparation) comprising: a) a set of about 1,000 arrayed clones from each of the five non-normalized libraries of brain (CT0s), heart (CS0s), kidney (CU0s), aorta (CW0s), and placenta (CX0s). The resulting pool of approximately 5,000 clones represented about 33.3% of the final driver population. A set of about 2,000 arrayed clones from each of the five normalized libraries of brain (CT0), heart (CS0), kidney (CU0), aorta (CW0), and placenta (CX0). The resulting pool of about 10,000 clones represented about 66.6% of the final driver population.

BASE COUNT 160 a 149 c 113 g 221 t
ORIGIN
TAG-TISSUE=rat kidney pool
TAG_SEQ=CAAGACTGTC"

alignment_scores:
Quality: 511.50 Length: 118
Ratio: 4.608 Gaps: 1
Percent Similarity: 94.068 Percent Identity: 86.441

alignment_block:
US-09-327-750d-35 x BI289546/rev ..

Align seg 1/1 to reverse of: BI289546 from: 1 to: 643

```

1 MetAlaSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
|||||
601 ATGGCGCTCCAAAGTCAACAAGTCATACCTGGATCTCAGTGTGGGAAAGA 552
|||||
16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluLup 33
|||||
551 CAAGAAAAACAAAAAGGTGGGAG.CCCTCCAAACAAGTGAACAGAAAT 503
|||||
33 rHisHisLeuGluValGluAsnLysLysProGlyGlyAsnValArg 49
|||||
502 CCCACCATCTGGAAGAGGTTGAAAAACAAGAGCCTGGGGCAATGTGAGG 453
|||||
50 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 66
|||||
452 AGGAAAGTCAGGCGACTGTGCTTAACCTTTCTATGGGCCATACCTAATAG 403
|||||
66 gHisValAspArgAsnGluGlyGluAspValGlyArgPheValValG 83
|||||
402 GCATGTTGATCATCAGTGAAGGGGAGGAGGTGGAGATTCCTAGGGC 353
|||||
83 InGlyThrGluValLysArgLysThrThrGluGlnValArgProTyr 99
|||||
352 AGGTGATGGAAGCAAGAGAAAGTCTAAGGAGCAACAGATGAGGCGCTTAC 303

```

```

KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 317)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:216854
Seq primer: mob.REGA+ET
High quality sequence stop: 291.
Location/Qualifiers
1..317
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:335454"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/notes="vector: pT7T3D (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(df) primer 15'
TGTACCAATCTGAAGTGGAGCGCGGCATTTTTTTTTTTT 3'[,
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."
```

```

BASE COUNT 108 a 63 c 87 g 59 t
ORIGIN
```

```

alignment_scores:
Quality: 541.00 Length: 101
Ratio: 5.410 Gaps: 0
Percent Similarity: 99.010 Percent Identity: 98.020
```

```

alignment_block:
US-09-327-750D-35 x W17712 ..
```

```

Align seg 1/1 to: W17712 from: 1 to: 317
```

```

17 LysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluPr 33
|||||
1 AAAAAAGACAAAAGAGTGGGAGGCGCTCCAAACAAGTGAAGAAGACC 50
|||||
33 oHisHisLeuGluGluValGluAsnLysProGlyGlyAsnValArgA 50
|||||
51 CCACCATCTGGAGAGAGTTGAAAACAAGAGCGCTGGGGAAATGTCGGA 100
|||||
50 rgLysValArgArgLeuValProAsnPhelLeuTrpAlaIleProAsnArg 66
|||||
101 CGAAAGTCAGGCGACTGTGTGCTTAACCTTCTCGGCCCATACCAATAGG 150
|||||
67 HisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValG1 83
|||||
151 CATGTTGATCGCAATGAAGGGGAGAGATGTTGGGAGATTTGTAGTGCA 200
|||||
```

```

83 nGlyThrGluValLysArgLysThrThrGluGlnValArgProTyra 100
|||||
201 GGAACAGAGGTCAGAGAAAGACTACGAGCAGCAGGTGAGGCCTTACA 250
|||||
100 rgArgPheArgThrProGluProAspAsnHisTyrAspPheCysLeulle 116
|||||
251 GCGGTTTCGGAACCCGGAACCTGACAAATCATACGACTTTTGCCCTCAT 300
|||||
117 Pro 117
|||
301 CCT 303
```

```

seq_name: gb_est1:AI227867
```

```

seq_documentation_block: 568 bp mRNA EST 20-JAN-1999
LOCUS AI227867
DEFINITION EST224562 Normalized rat brain, Bento Soares Rattus sp. cDNA clone
RBRCM82 3' end, mRNA sequence.
```

```

ACCESSION AI227867
VERSION AI227867.1 GI:3811754
KEYWORDS EST
```

```

SOURCE Rattus sp.
ORGANISM Rattus sp.
```

```

REFERENCE 1 (bases 1 to 568)
AUTHORS Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
Kerlavage,A.R. and Adams,M.D.
Rat Genome Project: Generation of a Rat EST (EST) Catalog & Rat
Gene Index
```

```

JOURNAL Unpublished (1998)
COMMENT Other ESTs: TC58086
Contact: Lee, NH
```

```

The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@igir.org
Seq primer: M13-21.
Location/Qualifiers
1..568
/organism="Rattus sp."
/db_xref="ATCC (inhost):2036891"
/db_xref="taxon:10118"
/clone="RBRCM82"
/clone_lib="Normalized rat brain, Bento Soares"
/notes="Organ: brain; Vector: pT7T3pac; Site_1: EcoRI;
Site_2: NotI"
```

```

FEATURES
source
```

```

BASE COUNT 140 a 143 c 108 g 177 t
ORIGIN
```

```

alignment_scores:
Quality: 539.50 Length: 118
Ratio: 4.860 Gaps: 1
Percent Similarity: 94.068 Percent Identity: 87.288
```

```

alignment_block:
US-09-327-750D-35 x AI227867/rev ..
```

```

Align seg 1/1 to reverse of: AI227867 from: 1 to: 568
```

```

1 MetaLysSerLys...LysGlnValIleLeuAspLeuThrValGluLysAs 16
|||||
528 ATGGCGTCCAAAAGTCAACAAGTCATACCTGATCTCACTGTGGAGAAGA 479
|||||
16 pLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluP 33
|||||
478 CAAGAAAAACAAAAGGTGGGAGGCGCTCCAAACAAGTGAAGAAGAAAT 429
|||||
33 roHisHisLeuGluGluValGluAsnLysProGlyGlyAsnValArg 49
|||||
```


OM of: US-09-327-750D-35 to: EST:* out_format : pfs

Date: Mar 11, 2002 2:16 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODE=framet_p2n.model -DEV=xlh
-O/cgnt2_1/USPTO.spool/US09327750/runat_11032002_101153_20308/app_query.fasta_1.1472
-DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.050 -LOOPEL=0.000 -LOOPEXT=0.000 -GAPEXT=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -EGAPEXT=6.000
-GAPEXT=7.000 -YGAPEXT=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=2000000000 -USER=US09327750 -CGN1_1_5654
-NCPU=6 -TCPUP=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-35

Query length: 117

Database: EST:*

Database sequences: 11351937

Database length: 1077921985

Search time (sec): 4085.940000

score_list:

Sequence	Strd	Orig	EScore	Len	Documentation
gb_btc:AK010400	+	611.50	1049.04	789	AK010400 Mus musculus ES cells
gb_est2:BG083261	+	611.50	1049.02	792	BG083261 H308C08-5 NIA Mouse 1
gb_est1:AA743525	+	605.00	1041.70	458	AA743525 vg78A09.r1 Barstead MF
gb_est2:BG070341	+	575.50	987.47	747	BG070341 H308C08-3 NIA Mouse 1
gb_est1:BF723075	+	568.50	977.99	519	BF723075 mab27e12.y1 Soares NME
gb_est1:BE333774	+	544.50	937.49	463	BE333774 mab27h06.y1 Soares NME
gb_est2:W17112	+	541.00	934.14	317	W17112 mb77e04.r1 Soares mouse
gb_est1:AI227867	+	539.50	927.44	568	AI227867 EST24552 Normalized
gb_est2:BI289946	+	511.50	878.38	643	BI289946 UI-R-DKO-cff-c-12-0-UI
gb_est1:AA117429	+	503.50	859.71	312	AA117429 mm21e12.r1 Bedington
gb_est2:W35893	+	484.50	834.09	472	W35893 mc5309.r1 Soares mouse
gb_est2:BF722027	+	479.00	828.16	286	BF722027 mab27e12.x1 Soares NME
gb_est2:W54487	+	469.50	808.38	465	W54487 md09c11.r1 Soares mouse
gb_est1:AL118340	+	463.50	798.74	422	AL118340 y911b26 Bedington mc
gb_est1:AW251984	+	459.00	789.25	540	AW251984 UI-R-BJ0-adn-b-06-0-UI
gb_est2:W83403	+	426.50	734.87	433	W83403 mf58d02.r1 Soares mouse
gb_est1:AI012535	+	373.00	642.62	444	AI012535 EST206986 Normalized
gb_est1:AV137854	+	368.00	636.78	300	AV137854 AV137854 Mus musculus
gb_est1:AA516739	+	365.00	631.85	290	AA516739 vh84c04.r1 Knowles Sol
gb_est2:W61422	+	363.50	630.70	237	W61422 md9g02.r1 Soares mouse
gb_est1:AL517301	+	355.50	609.93	639	AL517301 AL517301 LTI_NFL006_1
gb_est2:BE883414	+	355.50	609.07	722	BE883414 601509167F1 NIH_MGC_71
gb_est1:AL544124	+	355.50	607.95	845	AL544124 AL544124 LTI_NFL006_1
gb_est1:AU117899	+	355.50	607.79	865	AU117899 AU117899 HEMBAI Homo
gb_est2:BG122197	+	355.50	607.78	866	BG122197 602349748F1 NIH_MGC_90
gb_est1:AL545120	+	355.50	607.47	905	AL545120 AL545120 LTI_NFL006_1
gb_est2:BG709427	+	355.50	607.29	929	BG709427 602673481F1 NIH_MGC_96
gb_est2:BG035675	+	355.50	606.94	975	BG035675 602325745F1 NIH_MGC_90
gb_est1:AL560650	+	355.50	606.89	983	AL560650 AL560650 LTI_NFL010_1
gb_est1:AL539004	+	355.50	606.61	1023	AL539004 AL539004 LTI_FLO13_1
gb_est1:AL549875	+	355.50	606.46	1044	AL549875 AL549875 LTI_NFL006_1
gb_est2:BF930215	+	351.00	602.88	579	BF930215 IL5-NT0227-111200-319
gb_est1:AL534630	+	350.50	602.18	566	AL534630 AL534630 LTI_FLO13_1
gb_est2:AL523320	+	349.00	595.28	1043	AL523320 AL523320 LTI_NFL003_1
gb_est1:AW385267	+	347.50	597.43	534	AW385267 RC0-LT0001-261119-011
gb_est2:BG527431	+	347.50	594.82	772	BG527431 602557227F1 NIH_MGC_59
gb_est2:BG009563	+	346.50	595.11	581	BG009563 QV1-GN0319-021200-526
gb_est1:AL582060	+	346.00	590.16	1037	AL582060 AL582060 LTI_NFL010_1
gb_est2:BI040887	+	344.50	592.88	489	BI040887 QV3-NT0278-120201-502

gb_est1:AL573904 - 344.50 586.65 1.6e-23 1182
gb_est2:BI040889 - 340.50 584.98 2.0e-23 565
gb_est1:AW934883 + 338.50 580.35 3.7e-23 668
gb_est1:AW934905 + 338.50 580.18 3.8e-23 685
gb_est1:AW934769 + 338.00 579.56 4.1e-23 662
seq_name: gb_htc:AK010400

seq_documentation_block:
LOCUS AK010400 789 bp mRNA HTC 05-JUL-2001
DEFINITION Mus musculus ES cells cDNA, RIKEN full-length enriched library,
clone:2410004M13, full insert sequence.

ACCESSION AK010400
VERSION AK010400.1 GI:12845816
KEYWORDS CAP trapper.

SOURCE Mus musculus (strain:C57BL/6J) ES cells cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library
clone:2410004M13.

ORGANISM Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 789)

Carninci,P. and Hayashizaki,Y.

High-efficiency full-length cDNA cloning

Methods in enzymology. 303, 19-44 (1999)

99279253

PUBMED 10349636

2 (bases 1 to 789)

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome research. 10 (10), 1617-1630 (2000)

20493374

3 (bases 1 to 789)

Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P.,

Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,

Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,

Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,

Fujiwaka,S., Inoue,K., Tozawa,Y., Izawa,M., Ohara,E., Watahiki,M.,

Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuo,S., Kawai,J.,

Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome research. 10 (11), 1757-1771 (2000)

20530913

4 (bases 1 to 789)

Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,

Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,

Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F.,

Inotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,

Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,

Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,

Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,

Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,

Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takehashi,F.,

Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,

Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

Direct Submission

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,

URL:http://genome-gsc.riken.go.jp/, Tel:81-45-503-9222,

Fax:81-45-503-9216)

TITLE

JOURNAL

REFERENCE

5 (bases 1 to 789)

Nature 409, 685-690 (2001)

Functional annotation of a full-length mouse cDNA collection

FANTOM Consortium.

The RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium.

5 (bases 1 to 789)

Nature 409, 685-690 (2001)

Functional annotation of a full-length mouse cDNA collection

FANTOM Consortium.

The RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium.

Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10688 row: p column: 05
 High quality sequence stop: 712.

FEATURES

Location/Qualifiers

1..734

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:479476"

/clone_lib="NIH_MGC_96"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

187 a 174 c 214 g 159 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 646.00 Length: 125
 Ratio: 5.252 Gaps: 0
 Percent Similarity: 98.400 Percent Identity: 97.600

alignment_block:

US-09-327-750D-32 x BG715659 ..

Align seg 1/1 to: BG715659 from: 1 to: 734

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 206 ATGGAGTCCAAAGAGAAACGAGCAGTAACAGTCTCAGCATGGAATATGC 255
 17 aAsnGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProL 34
 256 CAACCAAGAAATGAAGGAAAGGAGCAAGTTGCTAATAAAGGGGAGCCCT 305
 34 euAlaLeuProLeuAspAlaGlyGlyTyrCysValProArgGlyAsnArg 50
 306 TGGCCCTCCCTTTGGATGCTGGTGAATACATGCTGCTGCTAGAGGAATCGT 355
 51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTyrAspIleMe 67
 356 AGCGCGTTCGCGCTTAGCGAGCCATCCTGCAGTATAGATGGATATGAT 405
 67 tHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
 406 GCATAGCTTTGGAGACCAACAGCAGGAGGATGAGAGAGAGATATGAAA 455
 84 rgileGlyGluGluValArgGlnLeuMetGluLysLysLeuArgGlyGln 100
 456 GGATTTGGGAGGAGGTGAGACACCTGATGGAAAGCTGAGGGAAAGCAG 505
 101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 117
 506 TTGAGTCATAGTCTGCGGGCAGTCAGCACTGACCCCTCCATCACCATGACCA 555
 117 sHisAspGluPheCysLeuMetPro 125
 556 TCATGATGATGTTTGGCTTATGCC 580

seq_name: gb_est2:BG709325

seq_documentation_block:

LOCUS BG709325 828 bp mRNA EST 07-MAY-2001
 DEFINITION 602673345F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4795882 5',
 mRNA sequence.

ACCESSION BG709325

VERSION BG709325.1 GI:13987550

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 828)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-re@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10679 row: j column: 11

High quality sequence stop: 798.

FEATURES

Location/Qualifiers

1..828

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4795882"

/clone_lib="NIH_MGC_96"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

232 a 179 c 228 g 189 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 637.00 Length: 126
 Ratio: 5.137 Gaps: 1
 Percent Similarity: 98.413 Percent Identity: 96.825

alignment_block:

US-09-327-750D-32 x BG709325 ..

Align seg 1/1 to: BG709325 from: 1 to: 828

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 176 ATGGAGTCCAAAGAGAAACGAGCAGTAACAGTCTCAGCATGGAATATGC 225
 17 aAsnGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProL 34
 226 CAACCAAGAAATGAAGAAAGGAGCAAGTTGCTAATAAAGGGGAGCCCT 275
 34 euAlaLeuProLeuAspAlaGlyGlyTyrCysValProArgGlyAsnArg 50
 276 TGGCCCTCCCTTTGGATGCTGCTGAATACATGCTGCTGCTAGAGGAATCGT 325


```
7111 CTAGAGGCCCGCCAGAAAAGAA.....AAGCAATACCACTAAAGTCTTT 7154
17 aasnGlnGluasnGluGlnValAlaAsnLysGlyGluProL 34
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
7155 AGAAGAAAGAACTAGAGTCAGAGAAACTCTAGAAATGGGGTCTCG 7204
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
34 euAlaLeuProLeu.....AspAlaGlyGlu 42
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
7205 TACTATCTGAGCTTTAGGAAAGACACACAGAGGATCAAGAA 7254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
43 TyrCysValProArgGlyAsnArgArgPhe..... 53
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
7255 TTAATGCTCTCTAAAGTACACTAAAGAGATTTTCATCTCTAGGAAAGGA 7304
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
54 .....ProValArgGlnProIleLeuGln..... 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
7305 AAGTCAGAAGTAGTGAGGCTTCAAAGAGGGGAACCTAGAAATCATGGA 7354
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 .....TyrArgTTPaspIleMetHisArgLeuGlyGluProGlnAlaArg 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
7355 CAGCTTTTAAAGAGGAGACCAACCCACTGGGATTTCCAGGAGCTGAG 7404
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
76 ..... 76
7405 GACCAGATGCTTGAGAGACTGCTAGAGAAAGAGGATCAGAGCTTCCCAAG 7454
77 .....MetArgGluGluA 81
7455 GTCTCCAGAGGAAGAGGACAGGAGGCATGTAGACCTCTGCAGAAAGAGA 7504
81 snMetGluArgIleGly.....GluGluValArgGlnLeuMetGluLys 95
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
7505 ATCAGGAACCATAGGTATGAGAGAGCAGAGGCCAGATACCTTGAGAGA 7554
96 LeuArgGluLysGlnLeuSerHisSerLeuArgAla 107
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
7555 CTGATAGAAAAGAGAGTCAAGGATCCCTGAGGTCT 7590

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-306-691B-23
```

```
seq_documentation_block:
; Sequence 23, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorana & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306.691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Moraco Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: NO. 5734039e
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2301 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-306-691B-23

alignment_scores:
   Quality: 84.00      Length: 102
   Ratio: 1.355      Gaps: 4
   Percent Similarity: 60.784      Percent Identity: 32.353

alignment_block:
US-09-327-750D-32 x US-08-306-691B-23 ..
Align seg 1/1 to: US-08-306-691B-23 from: 1 to: 2301

2 GluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAs 18
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
77 GAGGAGCAGGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
18 nGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProLeuA 35
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
127 GCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 176
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
35 laLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArgArg 51
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
177 CAAGCCGAG.....CGGAGGAGG 195
52 ArgPheProValArgGlnProIleLeu.....GlnTyrArgTrpAspI 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
196 CAGGAACCGGAGCGGAGGAGGAGTAGCTGGGTGGGACCATGCTGGCATCA 245
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
66 eMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMet 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
246 CCACCATCGAGCGGTGAAGCGCAAGATCCAGGTTCGACAGCAGCAGGCA 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
83 GluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArgGluL 99
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
296 GATGATGCAGAGGAGCGGAGCTGAGCGCCTCCAGCCGAGAACTTGAGG 345
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
99 ys 99
   : :
346 AA 347

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US93-06251-78

seq_documentation_block:
; Sequence 78, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
```

```

; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2301 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-78

alignment_scores:
  Quality: 84.00      Length: 102
  Ratio: 1.355       Gaps: 4
Percent Similarity: 60.784      Percent Identity: 32.353

alignment_block:
US-09-327-750D-32 x PCT-US93-06251-78 ..

Align seg 1/1 to: PCT-US93-06251-78 from: 1 to: 2301

2 GluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAs 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
77 GAGGAGCAGGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 126

18 nGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProLeuA 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
127 GCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 176

35 laLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArgArg 51
|| |||
177 CRAGCGCGAG.....CGGAGGAGG 195

52 ArgPheProValArgGlnProLeu.....GlnTyrArgTrpAspIle 66
::: ||| |||:|||||:|||||:|||||:|||||:|||||:|||||
196 CAGGAACCGGAGCGCGAGCAGTAGCTGGTGGCCACCACCATGGCTGATCA 245

66 eMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMet 82
||||| |||||| ||| |||:|||||:|||||:|||||:|||||:|||||
246 CCACCATCGAGCGGTGAACCGCACGATCCAGGTCTGCAGCAGCAGGCA 295

83 GluArgIleGlyGluGluValArgGlnLeuMet.GluLysLeuArgGluL 99
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
296 GATGATGCAGAGGAGCGAGCTGAGCGCCCTCCAGCGAGAGTGTGAGGAGA 345

99 ys 99
||
346 AA 347

```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-872-644-52

```

seq_documentation_block:
; Sequence 52, Application US/07872644
; Patent No. 5389527
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

```

```

; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/872,644
; FILING DATE: 19920420
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5389527and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2077 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1693
US-07-872-644-52

alignment_scores:
  Quality: 80.50      Length: 126
  Ratio: 1.118       Gaps: 6
Percent Similarity: 57.143      Percent Identity: 24.603

alignment_block:
US-09-327-750D-32 x US-07-872-644-52 ..

Align seg 1/1 to: US-07-872-644-52 from: 1 to: 2077

4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnG 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
959 CAGAGCGGTTCGAGTTTGAATAGCATCAGCTCGTCAGATCCCAAGCGATC 1008

20 u.....AsnGluGluLysGluGlnValAlaAsnLys. 30
::: |||:|||||:|||||:|||||:|||||:|||||:|||||
1009 AGGTGTCAAGACCTCTGTTTCAGAGGAGAGTCCCGCATCAACAATCTG 1058

31 GlyGluProLeu.....AlaLeuProLeuAspAlaGlyGluTyrCy 44
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1059 TCATCTCCGTTGACTATAAGAGCTTTAAAGCTACTTTGGACGGAAGTGTG 1108

44 sVal.....ProArgGlyAsnArgArgArgPheP 54
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1109 CACATCAATCGGAGAGATGGAGGCGCAAGGTACCCAAAGAGGAGGAGGC 1158

54 roValArgGlnProIleLeuGlnTyrArgTrpAsp.....IleMetHis 68
|| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1159 CAAGAAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1208

69 ArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArgI 85
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1209 AGGAAATGGAAGCAAAAGCGAGGCTGAAGAGGCGCATCTGCGCAAGCT 1258

```

Align seq 1/1 to: US-08-297-494-52 from: 1 to: 2077

[illegible]


```

; NAME: No. 5800987and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2077 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1693
; US-08-455-525-52

alignment_scores:
    Quality: 80.50      Length: 126
    Ratio: 1.118      Gaps: 6
    Percent Similarity: 57.143      Percent Identity: 24.603

alignment_block:
US-09-327-750D-32 x US-08-455-525-52

Align seg 1/1 to: US-08-455-525-52 from: 1 to: 2077

4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnG1 20
      ::::::::::::::::::::||| ::::::::::::::::::::
959 CAGAGCGGTTCGAGTTGTAATAGCATCAGCTCGTCAGATGCCAAGGCATC 1008
      ::::::::::::::::::::||| ::::::::::::::::::::
20 u.....AsnGluGluLysGluGlnValAlaAsnLys. 30
      ::::::::::::::::::::||| ::::::::::::::::::::
1009 AGGTGTCAAGACCTCTGGTTTCAGAGGGAAGTGCCTCGCATCAACATTCG 1058
      ::::::::::::::::::::||| ::::::::::::::::::::
31 GlyGluProLeu.....AlaLeuProLeuAspAlaGlyGluTyrCy 44
      ::::::::::::::::::::||| ::::::::::::::::::::
1059 TCATCTCCGTTGACTATAGAGCTTTAAAGCTACTTGGAGCGGAAGTGGTG 1108
      ::::::::::::::::::::||| ::::::::::::::::::::
44 sVal.....ProArgGlyAsnArgArgPhep 54
      ::::::::::::::::::::||| ::::::::::::::::::::
1109 CACATCAATCGGAGAGATGGAGGCCAAGGTACCCAAAGAGAGAGAAGC 1158
      ::::::::::::::::::::||| ::::::::::::::::::::
54 roValArgGlnProLeuLeuGlnTyrArgTrpAsp.....IleMetHis 68
      ::::::::::::::::::::||| ::::::::::::::::::::
1159 CAAGAAGGAAGCAGAGAGAAAGGCTCGCTGCCCGCAGAGAGCAGCAAA 1208
      ::::::::::::::::::::||| ::::::::::::::::::::
69 ArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArgI1 85
      ::::::::::::::::::::||| ::::::::::::::::::::
1209 AGGAATGGAGCCAAAGCCAGGCTGAGAGAGGCGCATCTGGCAAGACT 1258
      ::::::::::::::::::::||| ::::::::::::::::::::
85 eGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGlnLeuS 102
      ::::::::::::::::::::||| ::::::::::::::::::::
1259 GAGA.....AAAAGACGCTGTGGAGAACTAAGAACATCAAGTCAATGG 1299
      ::::::::::::::::::::||| ::::::::::::::::::::
102 erHisSerLeuArgAlaValSerThr 110
      ::::::::::::::::::::||| ::::::::::::::::::::
1300 AACACGGGCAACAAAGTGACAACC 1325
      ::::::::::::::::::::||| ::::::::::::::::::::

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-139-491-52

seq_documentation_block:
; Sequence 52, Application US/09139491
; Patent No. 6015677
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
; APPLICANT: Sonnenburg, William K.
; TITLE OF INVENTION: DNA Encoding Mammalian
;
; TITLE OF INVENTION: Phosphodiesterases
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/139,491
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,525
; FILING DATE: 31-MAY-1995
; APPLICATION NUMBER: 08/297,494
; FILING DATE:
; APPLICATION NUMBER: US 07/688,356
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6015677and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2077 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1693
; US-09-139-491-52

alignment_scores:
    Quality: 80.50      Length: 126
    Ratio: 1.118      Gaps: 6
    Percent Similarity: 57.143      Percent Identity: 24.603

alignment_block:
US-09-327-750D-32 x US-09-139-491-52

Align seg 1/1 to: US-09-139-491-52 from: 1 to: 2077

4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnG1 20
      ::::::::::::::::::::||| ::::::::::::::::::::
959 CAGAGCGGTTCGAGTTGTAATAGCATCAGCTCGTCAGATGCCAAGGCATC 1008
      ::::::::::::::::::::||| ::::::::::::::::::::
20 u.....AsnGluGluLysGluGlnValAlaAsnLys. 30
      ::::::::::::::::::::||| ::::::::::::::::::::
1009 AGGTGTCAAGACCTCTGGTTTCAGAGGGAAGTGCCTCGCATCAACATTCG 1058
      ::::::::::::::::::::||| ::::::::::::::::::::
31 GlyGluProLeu.....AlaLeuProLeuAspAlaGlyGluTyrCy 44
      ::::::::::::::::::::||| ::::::::::::::::::::
1059 TCATCTCCGTTGACTATAGAGCTTTAAAGCTACTTGGAGCGGAAGTGGTG 1108
      ::::::::::::::::::::||| ::::::::::::::::::::
44 sVal.....ProArgGlyAsnArgArgPhep 54
      ::::::::::::::::::::||| ::::::::::::::::::::
1109 CACATCAATCGGAGAGATGGAGGCCAAGGTACCCAAAGAGAGAGAAGC 1158
      ::::::::::::::::::::||| ::::::::::::::::::::
54 roValArgGlnProLeuLeuGlnTyrArgTrpAsp.....IleMetHis 68
      ::::::::::::::::::::||| ::::::::::::::::::::
1159 CAAGAAGGAAGCAGAGAGAAAGGCTCGCTGCCCGCAGAGAGCAGCAAA 1208
      ::::::::::::::::::::||| ::::::::::::::::::::
69 ArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArgI1 85
      ::::::::::::::::::::||| ::::::::::::::::::::
1209 AGGAATGGAGCCAAAGCCAGGCTGAGAGAGGCGCATCTGGCAAGACT 1258
      ::::::::::::::::::::||| ::::::::::::::::::::
85 eGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGlnLeuS 102
      ::::::::::::::::::::||| ::::::::::::::::::::
1259 GAGA.....AAAAGACGCTGTGGAGAACTAAGAACATCAAGTCAATGG 1299
      ::::::::::::::::::::||| ::::::::::::::::::::
102 erHisSerLeuArgAlaValSerThr 110
      ::::::::::::::::::::||| ::::::::::::::::::::
1300 AACACGGGCAACAAAGTGACAACC 1325
      ::::::::::::::::::::||| ::::::::::::::::::::

```

```

54 roValArgGlnProIleLeuGlnTyrArgTrpAsp.....lleMethHis 68
|||:::
1159 CAAGAAGGAGCAGAGGAAAAGCTCGCTGGCCGCGAGAGGACGACAA 1208
|||:::
69 ArgLeuGlyCluProGlnAlaArgMetArgGluGluAsnMetGluArgII 85
|||:::
1209 AGGAAATGGAAGCAAAAGCAGGCTGAAGAGGCGCATCTGGCAAAGCT 1258
|||:::
85 eGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGlnLeuS 102
:||||:::
1259 GAGA.....AAAGAGCTGTGGAGAAACTAAGAATCAAGTCAATGG 1299
:||||:::
102 erHisSerLeuArgAlaValSerThr 110
|||:::
1300 AACACGGGCAACAAAGAGTGAACACC 1325
|||:::

```

seq_name: /cgn2_6/ptodata/2/1na/PCTUS_COMB.seq:PCT-US92-03222-52

```

seq_documentation_block:
: Sequence 52, Application PC/TUS9203222
: GENERAL INFORMATION:
: APPLICANT: Beavo, Joseph A.
: APPLICANT: Bentley, Kelley
: APPLICANT: Charbonneau, Harry
: APPLICANT: Sonnenburg, William K.
: TITLE OF INVENTION: DNA Encoding Mammalian
: TITLE OF INVENTION: Phosphodiesterases
: NUMBER OF SEQUENCES: 58
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
: ADDRESSEE: Bicknell
: STREET: Two First National Plaza, 20 South Clark
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60603
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US92/03222
: FILING DATE: 19920420
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/688,356
: FILING DATE: 04-APR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Noland, Greta E.
: REGISTRATION NUMBER: 35,302
: REFERENCE/DOCKET NUMBER: 27866/30822
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 346-5750
: TELEFAX: (312) 984-9740
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 52:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2077 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 2..1693
PCT-US92-03222-52

```

alignment_scores: Quality: 80.50 Length: 126

```

Ratio: 1.118 Gaps: 6
Percent Similarity: 57.143 Percent Identity: 24.603
alignment_block:
US-09-327-750d-32 x PCT-US92-03222-52
Align seg 1/1 to: PCT-US92-03222-52 from: 1 to: 2077
4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnG1 20
:||||:::
959 CAGAGCGGTTCGAGTTTGAATAGCATCAGCTCGTCAGATGCCAAGCATC 1008
:||||:::
20 u.....AsnGluGluLysGluGlnValAlaAsnLys. 30
:||||:::
1009 AGGTGTCAAGACCTCTGGTTTCAGAGGGAAGTGCCTCCGATCAACAATCTG 1058
:||||:::
31 GlyGluProLeu.....AlaLeuProLeuAspAlaGlyGluTyrCy 44
:||||:::
1059 TCATCTCCGTTGACTATAGAAGCTTTAAAGCTACTTTGGACGGAAGTCGTG 1108
:||||:::
44 sVal.....ProArgGlyAsnArgArgPhep 54
:||||:::
1109 CACATCAATCGGAGAGATGGAGGCCAAGGTACCCAAAGAGAGGAAGGC 1158
:||||:::
54 roValArgGlnProIleLeuGlnTyrArgTrpAsp.....lleMethHis 68
|||:::
1159 CAAGAAGGAGCAGAGGAAAAGGCTCGCTGGCCGCGAGAGGACGACAA 1208
|||:::
69 ArgLeuGlyCluProGlnAlaArgMetArgGluGluAsnMetGluArgII 85
|||:::
1209 AGGAAATGGAAGCAAAAGCAGGCTGAAGAGGCGCATCTGGCAAAGCT 1258
|||:::
85 eGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGlnLeuS 102
:||||:::
1259 GAGA.....AAAGAGCTGTGGAGAAACTAAGAATCAAGTCAATGG 1299
:||||:::
102 erHisSerLeuArgAlaValSerThr 110
|||:::
1300 AACACGGGCAACAAAGAGTGAACACC 1325
|||:::

```

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-07-872-644-50

```

seq_documentation_block:
: Sequence 50, Application US/07872644
: Patent No. 5389527
: GENERAL INFORMATION:
: APPLICANT: Beavo, Joseph A.
: APPLICANT: Bentley, Kelley
: APPLICANT: Charbonneau, Harry
: APPLICANT: Sonnenburg, William K.
: TITLE OF INVENTION: DNA Encoding Mammalian
: TITLE OF INVENTION: Phosphodiesterases
: NUMBER OF SEQUENCES: 58
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
: ADDRESSEE: Bicknell
: STREET: Two First National Plaza, 20 South Clark
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60603
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/872,644
: FILING DATE: 19920420
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/688,356

```

```
; FILING DATE: 04-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5389527and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2693 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 176..2077
; US-07-872-644-50

alignment_scores:
  Quality: 80.50      Length: 126
  Ratio: 1.118      Gaps: 6
  Percent Similarity: 57.143      Percent Identity: 24.603

alignment_block:
US-09-327-750D-32 x US-07-872-644-50 ..

Align seg 1/1 to: US-07-872-644-50 from: 1 to: 2693

4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnG1 20
1568 CAGAGCGGTCGAGTTGATAGATCAGCTCTGATGCGGAGGAGGATC 1617
20 u.....AsnGluLysGluGlnValAlaAsnLys. 30
1618 AGGTGTCAAGACCTCTGTTTCAGAGGGAAGTCCCGCATCAACAATTCG 1667
31 GlyGluProLeu.....AlaLeuProLeuAspAlaGlyGluTyrCy 44
1668 TCATCTCCGTTGACTATAAGAGCTTTAAAGCTACTTTGGACGGAAGTGTG 1717
44 sVal.....ProArgGlyAsnArgArgArgPhep 54
1718 CACATCAATCGGAGAGATGGAGGCGCAAGGTACCCAGAGAGAGAGGC 1767
54 roValArgGlnProLeuGlnTyrArgTrpAsp.....IleMetHis 68
1768 CAAGAAGGAAGCAGAGAGAAAGGCTCGCTCGCGCAGAGGAGCAGCAAA 1817
69 ArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArgI1 85
1818 AGGAATGAAGCCAAAGCCAGGCTTGAAGAGCGCATCTCGCAAGCT 1867
85 eGlyGluGluValArgGluMetGluLysLeuArgGluLysGlnLeuS 102
1868 GAGA.....AAAGACGCTCTGGAGAACTAAGATCAAGTCAATGG 1908
102 erHisSerLeuArgAlaValSerThr 110
1909 AACACGGGCAACAAAGTGACAACC 1934

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-297-494-50

seq_documentation_block:
; Sequence 50, Application US/08297494
; Patent No. 5580771
; GENERAL INFORMATION:
; APPLICANT: Beavo, Joseph A.
; APPLICANT: Bentley, Kelley
; APPLICANT: Charbonneau, Harry
```

[illegible]

seq_name: /cqn2_6/ptodata/2/ina/5A_COMB.seq:US-08-479-532-50

seq_documentation_block;
; Sequence 50, Application US/08479532
; Patent No. 5776752

; Patent No. 5776752
: GENERAL INFORMATION:

APPLICANT: Beavo, Joseph A.

APPLICANT: Bentley, Kelley

; APPLICANT: Charbonneau, Harry

APPLICANT: Sonnenburg, William K

;	TITLE OF INVENTION:	DNA ENCODING
;	TITLE OF INVENTION:	phosphodiester

NUMBER OF SEQUENCES: 58

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, & Riehn

ADDRESSEE: Bicknell
STREET: Two First National Plaza

STREET: TWO FIRST NATIONAL PL
STREET: Street
STREET: Street

CITY: Chicago

STATE: Illinois

```

;
; COUNTRY: USA
; ZIP: 60603

```

ZIP: 60603
COMPUTER READABLE FORM:

```
;
; CONT OF LER REMAINING FORM.
; MEDIUM TYPE: Floppy disk
```

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE:  PatentIn Release #1.
: CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/479,100

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

;; PRIOR APPLICATION DATA: 08/297-494

```

; APPLICATION NUMBER: 08/237,434
; FILING DATE:

```

100

THIS PAGE BLANK (USPTO)

OM of: US-09-327-750D-32 to: N_Geneseq_1101.* out_format : pfs
Date: Mar 11, 2002 3:42 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=framet_p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09327750/runat_11032002_101154_20362/app_query.fasta_1.1472
-DB=N_Geneseq_1101 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FCAPOP=6.000 -FCAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09327750.@CGL1_1330 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-327-750D-32
Query length: 125
Database: N_Geneseq_1101.*
Database sequences: 930621
Database length: 428662619
Search time (sec): 355.560000

score_list:	Strd Orig	zScore	EScore Len	Documentation
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143	+	639.00	1250.82	792
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH75810	+	602.50	1216.52	767
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH58581	+	602.50	1215.36	862
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH59611	+	602.50	1214.95	898
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03067	+	589.50	1188.82	858
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH03880	+	585.50	1183.24	862
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH03879	+	585.50	1182.71	698
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01089	+	562.50	1135.77	692
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH06183	+	207.50	414.96	421
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517	+	202.50	397.55	457
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13750	+	202.50	394.04	865
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH83548	+	202.50	393.00	1229
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH23529	+	201.00	394.18	1364
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH23528	+	185.50	362.21	891
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH23528	+	185.50	363.88	917
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH06184	+	184.50	367.25	700
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01005	+	151.50	298.13	451
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01004	+	136.50	268.39	532
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH06181	+	117.00	234.94	485
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01003	+	106.00	207.93	279
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH87147	+	100.50	204.30	401
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAH14053	+	89.00	160.28	1459
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH82586	+	88.50	171.55	426
/SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT:AAH28398	+	86.50	141.15	5946
/SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT:AAH70447	+	86.50	134.79	11236
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAH20445	+	84.00	145.51	2301
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH09303	+	84.00	145.51	2301
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH15730	+	81.00	135.10	3532
/SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT:AAH30188	+	80.50	139.38	2077
/SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:AAH83982	+	80.50	139.38	2077
/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAH67203	+	80.50	139.38	2077
/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAH51117	+	80.50	139.38	2077
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAH54766	+	80.50	139.38	2077
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAH48209	+	80.50	139.38	2077
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAH36161	+	80.50	139.38	2077
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH290385	+	80.50	136.79	2693
/SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:AAH83981	+	80.50	136.79	2693
/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAH67202	+	80.50	136.79	2693
/SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:AAH51116	+	80.50	136.79	2693

/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV54765	+	80.50	136.79	15.27
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV48208	+	80.50	136.79	15.27
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV36160	+	80.50	136.79	15.27
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ90384	+	80.50	136.79	15.27
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ20086	+	80.50	124.90	70.07
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143				
seq_documentation_block:				
ID	AAH45143 standard; cDNA; 792 BP.			
XX	AAH45143;			
AC	07-SEP-2001 (first entry)			
DT				
XX	Human brain expressed X-linked protein, hbex, coding sequence.			
DE				
XX	Human; brain expressed X-linked protein; cytostatic; auditory; nontropic;			
KW	hbex; dysembryoplasia; hereditary disease; cancer; tumour; deafness;			
KW	X-chromosome-binding mental retardation; lissencephalous disease; ss.			
XX	Homo sapiens.			
OS	WO200140286-A1.			
PN	07-JUN-2001.			
XX				
PD	27-NOV-2000; 2000WO-CN00502.			
XX				
PF	30-NOV-1999; 99CN-0124179.			
XX				
PR	(BIOR-) BIOROAD GENE DEV LTD SHANGHAI.			
PA				
XX	Mao Y, Xie Y;			
PI	WPI; 2001-397944/42.			
XX	P-PSDB; AAB99224.			
DR	Isolated human brain-expressed X-linked polypeptide used to diagnose			
PT	and treat of dysembryoplasia, hereditary diseases, cancer, tumor,			
PT	deafness and X-chromosome-binding mental retardation			
XX	Claim 5; Page 22; 30pp; Chinese.			
PS				
XX				
CC	The present sequence is the coding sequence for a human brain-expressed			
CC	X-linked protein (hbex). hbex and its coding sequence are useful in the			
CC	diagnosis and treatment of dysembryoplasia, hereditary diseases, cancer,			
CC	tumours, deafness, X-chromosome-binding mental retardation and			
CC	lissencephalous disease. hbex is also useful for screening mimics,			
CC	agonists, or inhibitors, and in peptide fingerprinting identification.			
CC	hbex coding sequence can be used as primers or probes, or in producing			
CC	gene chips or microarrays.			
XX	Sequence 792 BP; 214 A; 172 C; 219 G; 187 T; 0 other;			
SQ				
alignment_scores:				
Quality: 639.00		Length: 125		
Ratio: 5.238		Gaps: 0		
Percent Similarity: 97.600		Percent Identity: 96.800		
alignment_block:				
US-09-327-750D-32 x AAH45143				
Align seg 1/1 to: AAH45143 from: 1 to: 792				
1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17				
173 ATGGAGTCCNAGAGAACTAGCAGTAAACAGTCTCAGCATGGAATGC 222				
17 aAsnGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProL 34				
223 CAACCAAGAAATGAAGAAAAGGAGCAAGTTGCTAATAAAGGGAGGCCT 272				

34 euAlaLeuProLeuAspAlaGlyCysValProArgGlyAsnArg 50
|||||
273 TGCCCTCCCTTTGGATGCTGGTGAATACTGTGTGCTAGAGGAATCGT 322

51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIle 67
|||||
323 AGCGGTTCCGCTTAGCGAGCCCATCTGCAGTAGATGGATATGAT 372

67 tHisArgLeuGlyCysProGlnAlaArgMetArgGluGluAsnMetGlu 84
|||||
373 GCATAGCTTGGAGAACACAGGCAAGGATGAGAGAGAGAAATATGAAA 422

84 rGileGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
|||||
423 GGATGGGGAGGGGTGAGACAGCTGATGGAAGAAGCTGAGGGAAGACAG 472

101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 117
|||||
473 TTGAGTCATAGTCTGGGGCAGTCAGCACTGACCCCTCACCATGACCA 522

117 sHisAspGluPheCysLeuMetPro 125
|||||
523 TCATGATGAGTTTGGCTTATGCCC 547

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH75810

seq_documentation_block:
ID AAH75810 standard; cdna; 767 BP.
XX AC AAH75810;
XX
DT 17-OCT-2001 (first entry)
XX
DE Human X chromosome linked gene expression protein 14 coding sequence.
XX
KW Human; X chromosome linked gene expression protein 14; cancer;
KW HIV infection; cytostatic; anti-HIV; chromosome X; ss.
XX
OS Homo sapiens.
XX
PN CN1296969-A.
XX
PD 30-MAY-2001.
XX
PF 23-NOV-1999; 99CN-0124078.
XX
PR 23-NOV-1999; 99CN-0124078.
XX
PA (SHAN-) SHANGHAI BORONG GENE DEV CO LTD.
XX
PL Mao Y, Xie Y;
XX
DR WPI; 2001-483897/53.
DR P-PSDB; AAG66407.
XX
PT Polypeptide-human X chromosome linked gene expression protein 14 and
PT polynucleotide for coding said polypeptide -
XX
PS Claim 6; Page 24 (Disclosure); 31pp; Chinese.
XX

The present sequence is the coding sequence for human X chromosome linked gene expression protein 14. The protein and coding sequence are useful for treating diseases e.g. cancer and HIV infection.

Sequence 767 BP; 224 A; 143 C; 215 G; 185 T; 0 other;

alignment_scores:
Quality: 602.50 Length: 128
Ratio: 4.898 Gaps: 1
Percent Similarity: 96.094 Percent Identity: 89.062

alignment_block:
US-09-327-750D-32 x AAH75810 ..

Align seg 1/1 to: AAH75810 from: 1 to: 767

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
101 ATGGAGTCCAAAGAGGACGAGCGTTAAACAATCTCATCTGGAAAATGT 150

17 asnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
:|||||
151 CAACACAGGAAATGATGAAANAGATGAAAGGAGCAAGTTGCTTAATAAG 200

31 LysGluProLeuAlaLeuProLeuAspAlaGlyCysValProArg 47
|||||
201 GGGAGCCCTTGGCCCTTACCTTTGAATGTTAGTGAATACACTGTGCGCTAGA 250

48 GlyAsnArgArgPheProValArgGlnProIleLeuGlnTyrArgTr 64
|||||
251 GGAAACCGTAGCGGTTCCGCGTTAGGAGCCCATCTGCAGTATAGATG 300

64 pAspIleMetHisArgLeuGlyCysProGlnAlaArgMetArgGluGlu 81
|||||
301 GGACATAATGCATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGAGAGA 350

81 snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 97
|||||
351 ATATGGAAAGGATTTGGGAGGAGGTGAGACAGCTGATGGAAGAAGCTGAGG 400

98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProHis 114
|||||
401 GAAAGCAGCTTGAGTCATAGTCTGGGGCAGTCAGCACTGATCCCTCA 450

114 sHisAspHisHisAspGluPheCysLeuMetPro 125
|||||
451 CCATGACCATCAGTACGAGTTTGGCTTATGCCC 484

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI58581

seq_documentation_block:
ID AAI58581 standard; cdna; 862 BP.
XX AC AAI58581;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 784.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-brager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-052317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR P-PSDB; AAM39425.
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX Claim 1: SEQ ID NO 784; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 862 BP; 235 A; 183 C; 238 G; 206 T; 0 other;

alignment_scores:

Quality: 602.50 Length: 128
Ratio: 4.898 Gaps: 1
Percent Similarity: 96.094 Percent Identity: 89.062

alignment_block:

US-09-327-750d-32 x AAI58581 ..

Align seg 1/1 to: AAI58581 from: 1 to: 862

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
226 ATGGAGTCCAAAGAGAACGACGCGTTAAACAATCTCATCGTGGAAATGT 275
17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
276 CAACCAAGAAATGATGATAAAGATGAAAGGACGCAAGTGTCTAATAAG 325
31 lYGlupProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
326 GGGACCCCTTGGCCCTACCTTTGAATGTAGTGAATACTGTGTGCTAGA 375
48 GlyAsnArgArgPheProValArgGlnProIleLeuGlnTyrArgTr 64
376 GGAACCGTAGCGGTTCCGTTAGGAGACCCATCCCTGCAGTATAGATG 425
64 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 81
426 GCACATAATGCATAGCTTGGAGAGCCACAGGCAAGGATGACAGAGGAGA 475
81 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 97
476 ATATGGAAGAGGATGGGAGGAGGTGACACAGCTGATGAAAGCTGAGG 525
98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 114
526 GAAAGCAGTTCAGTCATAGTTTGGCGGCGAGTCAGCAGTGTGCCCTCA 575
114 shHisAspHisAspGluPheCysLeuMetPro 125
|||||

576 CCATGACCATCAGCATGAGTGTGCTTATGCCCC 609
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF59611
seq_documentation_block:
ID AAF59611 standard; cDNA; 898 BP.
XX
AC AAF59611;
XX
DT 24-APR-2001 (first entry)
XX
DE Human cell cycle and proliferation protein CCYPR-22 cDNA, SEQ ID NO:76.
XX
KW Cell cycle and proliferation protein; CCYPR; human; agonist;
KW antagonist; gene therapy; detection; gene therapy;
KW transgenic animal disease model; immune disorder;
KW developmental disorder; cell signalling disorder;
KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
KW arteriosclerosis; asthma; allergy; diabetes mellitus;
KW menstrual cycle disorder; bacterial infection; ss.
XX
OS Homo sapiens.
XX
WO200107471-A2.
XX
PD 01-FEB-2001.
XX
PF 21-JUL-2000; 2000WO-US19948.
XX
PR 21-JUL-1999; 99US-0145075.
PR 08-SEP-1999; 99US-0153129.
PR 10-NOV-1999; 99US-0164647.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
PI Azimzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;
XX
DR WPI: 2001-112727/12.
DR P-PSDB; AAB60474.
XX
PT Human cell cycle and proliferation proteins and polynucleotides are
PT used to treat, diagnose and prevent immune, developmental and cell
PT signaling disorders and cell proliferative disorders including cancer -
XX
PS Claim 5; Page 181-182; 205pp; English.
XX
CC Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human
CC cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.
CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
CC associated with decreased expression of functional CCYPR, while CCYPR
CC antagonists are used to treat diseases or conditions associated with
CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
CC that specifically bind to CCYPR, and in drug screening methods to
CC identify compounds that modulate the activity of CCYPR. CCYPR
CC nucleotides can be used to generate transgenic animal models of human
CC disease, and can be used in gene therapy in target cells with genetic
CC abnormalities with respect to the expression of CCYPR for the
CC treatment or prevention of a disorder associated with CCYPR.
CC Diseases which can be diagnosed, treated and prevented using CCYPR
CC proteins, nucleic acids, agonists or antagonists include immune,
CC developmental and cell signalling disorders, and cell proliferative
CC disorders including cancer. Specific examples of these disorders
CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
CC diabetes mellitus, disorders of the menstrual cycle and infections
CC caused by bacteria.
XX
SQ Sequence 898 BP; 250 A; 186 C; 251 G; 211 T; 0 other;

```
alignment_scores:
  Quality: 602.50      Length: 128
  Ratio: 4.898         Gaps: 1
  Percent Similarity: 96.094  Percent Identity: 89.062

alignment_block:
US-09-327-750D-32 x AAF59611
Align seg 1/1 to: AAF59611 from: 1 to: 898

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
254 ATGGAGTCCAAAGAGGAACGCGTTTAAACAATCTCATCGTGGAAAAATGT 303
17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
304 CRACCAAGAAATGATGAAAGAGATGAAAGGAGCAAGTTGCTAATAAG 353
31 lyGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
354 GGGAGGCCCTTGGCCCTACCTTTGAATGTTAGTGAATCTGCTGCCTAGA 403
48 GlyAsnArgArgPheProValArgGlnProIleLeuGlnTyrArgTr 64
404 GGAACCGTGGCGGTTCGCCCTTAGGCAGCCCATCTGCAGTATAGATG 453
64 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu 81
454 GGACATAATGCATAGCTTGGAGAGCCACAGCAAGGATGAGAGAGAGA 503
81 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 97
504 ATATGAAAGGATTGGGAGGAGGTGAGACAGCTGATGAGAAAGCTGAGG 553
98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 114
554 GAAAGCAGTTGAGTCATAGTCTGGGGCAGTCAGACACTGATCCCCCTCA 603
114 shisAspHisHisAspGluPheCysLeuMetPro 125
604 CCATGACCATCAGTATGAGTTTGGCTTATGCCC 637
```

seq_name: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI60367

```
seq_documentation_block:
ID AAI60367 standard; cdNA; 858 BP.
XX
AC AAI60367;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4356.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; SS.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0398042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
```

```
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang JT, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR P-PSDB; AAM41211.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PS Claim 1; SEQ ID NO 4356; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM4213).with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 858 BP; 236 A; 175 C; 241 G; 206 T; 0 other;
```

```
alignment_scores:
  Quality: 589.50      Length: 129
  Ratio: 4.793         Gaps: 2
  Percent Similarity: 95.349  Percent Identity: 88.372

alignment_block:
US-09-327-750D-32 x AAI60367
Align seg 1/1 to: AAI60367 from: 1 to: 858

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
215 ATGGAGTCCAAAGAGGAGGAGCGTTTAAACAATCTCATCGTGGAAAAATGT 264
17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
265 CAACCAAGGAAATATGATGAAAAAGATGAAAGGAGCAAGTTGCTAATAAG 314
31 lyGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
315 GGGAGCCCTTGGCCCTACCTTTGAATGTTAGTGAATCTGCTGCTAGAGA 364
48 GlyAsnArgArgPheProValArgGlnProIleLeuGlnTyrArgTr 64
365 GGAACCGCTAGCGGTTCCGCGTTAGGAGCCCATCTCGCAGTATAGATG 414
64 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu 81
415 GGACATAATGCTAGCTTGGAGAGCCACAGCAAGGATGAGAGAGAGA 464
81 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 97
465 ATATGAAAGGATTGGGAGGAGGTGAGACAGCTGATGAGAAAGCTGAGG 514
98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 114
```


CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors.
 XX
 SQ Sequence 698 BP; 187 A; 151 C; 210 G; 139 T; 11 other;

alignment_scores:
 Quality: 585.50 Length: 128
 Ratio: 4.920 Gaps: 1
 Percent Similarity: 92.969 Percent Identity: 86.719

alignment_block:

US-09-327-750D-32 x AAC03879 ..

Align seg 1/1 to: AAC03879 from: 1 to: 698

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 |||||
 243 ATGGAGTCCAAAGAGAACGAGCGTTAAACAATCTCATCGTGGAAAATGT 292
 |||||
 17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysC 31
 |||||
 293 CAACAGGAAATATGATGAAAAGATGAAAAGGAGCAAGTGWCTTAATAAAG 342
 |||||
 31 lYlGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
 |||||
 343 GGGAGCCCTTGGCCCTACCTTGTGATGTYRGTAATCTGTGTCCTAGA 392
 |||||
 48 GlyAsnArgArgPheProValArgGlnProIleLeuGlnTyrArgTr 64
 |||||
 393 GGAATCGTAGGGCGTTCGCGTGTAGGACGCCCATCTGCAGTATAGATG 442
 |||||
 64 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 81
 |||||
 443 GGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 492
 |||||
 81 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 97
 |||||
 493 ATATGGAAGGATTTGGGAGGAGCTGAGACAGCTGATGGAAGAGCTGAGG 542
 |||||
 98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 114
 |||||
 543 GAAAGACAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 592
 |||||
 114 sHisAspHisHisAspGluPheCysLeuMetPro 125
 |||||
 593 CCATGACCATCATGATGATGATGATGATGATGATGATGATGATGATGATG 626
 |||||

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAC10889

seq_documentation_block:

ID AAC10889 standard; cDNA; 692 BP.

XX

AC AAC10889;

XX

DT 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 14964.

DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping; ss.

XX

OS Homo sapiens.

XX

PN EP1033401-A2.

XX

PD 06-SEP-2000.

XX

PF 21-FEB-2000; 2000EP-0200610.

XX

PR 26-FEB-1999; 99US-0122487.

XX

PA (GEST) GENSET.

XX

PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX

DR WPI: 2000-500381/45.

XX

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX

PS Claim 1; SEQ ID 14964; 7lpp + CD-ROM; English.

XX

CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.

XX

SQ Sequence 692 BP; 184 A; 151 C; 199 G; 146 T; 12 other;

alignment_scores:

Quality: 562.50 Length: 125

Ratio: 4.849 Gaps: 1

Percent Similarity: 92.800 Percent Identity: 84.800

alignment_block:

US-09-327-750D-32 x AAC10889 ..

Align seg 1/1 to: AAC10889 from: 1 to: 692

4 LysGluLysArgAlaValAsnSerLeuSerMetGluAsnAlaAsnGlnGl 20

|||||

246 CAAAGAGAACGAGCGTTAAACAATCTCATCGTGGAAAATGTCAACCCAGGA 295

|||||

20 uAsn.....GluGluLysGluGlnValAlaAsnLysGlyGluProL 34

|||||

296 AAATGATGAAAAGATGAAAAGGAGCAAGTGWCTTAATAAGGGGAGCCCT 345

|||||

34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50

|||||

346 TGGCCCTACCTTGTGATGTYRGTAATCTGTGCTAGAGGAATCGT 395

|||||

51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67

|||||

396 AGCGCGTTCGCGCTTAGGCAGGCCCATCTGCAGTATAGATGGGATATGAT 445

|||||

67 thisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84

|||||

446 GCATAGGCTTGGAGAACCAAGGATGAGAGAGAGATATATGGAA 495

|||||

84 rgIleGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100

|||||

496 GGATTGGCGAGGAGGTGACAGAGCTGATCGAAAGCTGAGGAAAGACAG 545

|||||

101 LeuSerHisSerLeuArgAlaValSerThrAspProProHisHisAspHl 117

|||||

546 TTGAGTCATAGTCTGCGGGCAGTCAGCACTGACCCCTCACCATGACCA 595

117 shiAspGluPheCysLeuMetPro 125
|||||
596 TCATGATGAGTTTGCNNWATGCC 620

seq_name: /SDS2/gcgdata/geneseq/geneseq/NA2000.DAT: AAC06183

seq_documentation_block:

ID AAC06183 standard; cDNA; 421 BP.

AC AAC06183;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 10258.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 1; SEQ ID 10258; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.

XX Sequence 421 BP; 110 A; 103 C; 124 G; 82 T; 2 other;

alignment_scores:

Quality: 207.50 Length: 72
Ratio: 3.402 Gaps: 2
Percent Similarity: 84.722 Percent Identity: 65.278

alignment_block:

US-09-327-750D-32 x AAC06183 ..

Align seg 1/1 to: AAC06183 from: 1 to: 421

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
207 ATGGAGTCCAAAGAGGACGCGTTAAACAATCTCATCTCGTGGAAAATGT 256

17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
:|||||
257 CAACAGGAAATGATGAAAGATGAAAGAGGACGAGTGTGCTAATAAG 306
|||||
31 lyGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
|||||
307 GGGAGCCCTTGGCCCTTACCTTTGAATGTAGTGAATAGTGTGTGCTAGA 356
|||||
48 GlyAsnArgArgPheProValArgGlnProIleLeuGln.TyrArgT 64
|||||
357 GGAACCCCTAG.CGTTCCGCTTAGGCACACAGTATTAGAGGCCCGCT 405
|||||
64 rpAspIleMetHis 68
::: |||
406 GCCCAGTGACACAT 419

seq_name: /SDS2/gcgdata/geneseq/geneseq/NA2000.DAT: AAC06180

seq_documentation_block:

ID AAC06180 standard; cDNA; 457 BP.

XX AAC06180;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 10255.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 1; SEQ ID 10255; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.

XX Sequence 457 BP; 120 A; 110 C; 142 G; 84 T; 1 other;

alignment_scores:

Quality: 207.50 Length: 72
Ratio: 3.402 Gaps: 2

Percent Similarity: 84.722 Percent Identity: 65.278

alignment_block:

US-09-327-750D-32 x AAC06180 ..

Align seg 1/1 to: AAC06180 from: 1 to: 457

```
1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
243 ATGGAGTCCAAAGAGACGACGGTTAAACAATCTCATCGTGGAAATGT 292
|||||
17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
|||||
293 CAACACAGAAATGATGAAAAAGATGAAAGGACGCAAGTTGCTAATAAG 342
|||||
31 lYcLupProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
|||||
343 GGGAGCCCTTGGCCCTACCTTTGAATGTAGTCAATACTGTGTGCTAGA 392
|||||
48 GlyAsnArgArgArgPheProValArgGlnProIleLeuGln.TyrArgT 64
|||||
393 GGAACCCGTAG.CGGTTCCCGCTTAGGCACACAGTATTAGAGCCCGCT 441
|||||
64 rpAspIleMethIs 68
::: |||
442 GCCCAGTGACACAT 455
```

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517

seq_documentation_block:

ID AAH03517 standard; cDNA; 865 BP.

XX AC AAH03517;

DT 26-JUN-2001 (first entry)

XX DE Human cDNA clone (5'-primer) SEQ ID NO:352.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isodai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WI WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

XX PS Claim 1; SEQ ID 352; 2537pp + CD ROM; English.

XX CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

XX SQ Sequence 865 BP; 239 A; 177 C; 224 G; 221 T; 4 other;

alignment_scores:

Quality: 202.50 Length: 126
Ratio: 2.382 Gaps: 3
Percent Similarity: 67.460 Percent Identity: 38.889

alignment_block:

US-09-327-750D-32 x AAH03517 ..

Align seg 1/1 to: AAH03517 from: 1 to: 865

```
1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
191 ATGGAGTCCAAAGAGACGACCTACGGCAACAATCTCAACGGGGAATATC 240
|||||
17 aAsnGlnGluAsnGluLysGluGlnValAlaAsnLysGlyGluProL 34
|||||
241 CCAACAAGAAACGAAGGAGGAGGAGCAGGCCCCACGCAAGTGAAGAAG 290
|||||
34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
::: |||
291 AATCCCGCCATTGGGAGGGGTGAAGGCCAGAGCCTGGAGGAATATC 340
|||||
51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67
|||||
341 AGCGGGGCGGAGTTAGGCGACTTGTCCCTAATTTTCGATGGCCATACC 390
|||||
67 thisArg...LeuGlyGluProGlnAlaArgMetArgGluGluAsnMetG 83
|||||
391 TAATAGGCATATTGAGCACAATGAAGCGAGA.....GATGATGTAG 431
|||||
83 luArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLys 99
|||||
432 AAAGGTTTGTAGGCGCAGATGATGGAATCAAGAGAAAGACTAGGGAACAG 481
|||||
100 GlnLeuSerHisSerLeuArgAlaValSerThrAspProProHisHisAs 116
|||||
482 CAGATAGGCACATATATGCGCTTCCAAACTCTCTGAACCT.....GA 522
|||||
116 pHishHisAspGluPheCysLeuMetPro 125
|||||
523 CAACCATATTAGACTTTTCCTCATACCT 550
```

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13750

seq_documentation_block:

ID AAH13750 standard; cDNA; 1229 BP.

XX AC AAH13750;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA sequence SEQ ID NO:10656.

CC in cell differentiation (CDIFF). CDIFF polypeptides and agonists of
 CC these are used to treat a disease or condition associated with
 CC decreased expression of functional CDIFF. An antagonist of CDIFF is
 CC used to treat a disease or condition associated with over expression
 CC of functional CDIFF. CDIFF polypeptides may be used for the treatment,
 CC prevention and diagnosis of cell proliferative, developmental and
 CC neurological disorders, such as Alzheimer's disease, schizophrenic
 CC disorders, arteriosclerosis, cancer, atherosclerosis, diabetes mellitus
 CC and epilepsy. The CDIFF-4 sequence is homologous to Mus musculus
 CC REX-3. This sequence maps to chromosome 1 within the interval from
 CC 152.2 to 157.4 centimorgans, to chromosome 3 within the interval from
 CC 157.4 to 158.0 centimorgans, and to the X chromosome within the interval
 CC from 104.9 to 150.3 centimorgans.

XX Sequence 1364 BP; 411 A; 269 C; 322 G; 362 T; 0 other;

alignment_scores:
 Quality: 202.50 Length: 126
 Ratio: 2.382 Gaps: 3
 Percent Similarity: 67.460 Percent Identity: 38.889

alignment_block:

US-09-327-750D-32 x AAC85548 ..

Align seg 1/1 to: AAC85548 from: 1 to: 1364

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 250 ATGGAGTCCAAAGAGAGAACTAGCGCAAAACAATCTCAACGGGGAAATGC 299
 17 aAsnGlnGluAsnGluLysGluGlnValAlaAsnLysGlyGluProL 34
 1:|||||
 300 CCAACAGAAAGAAAGAGGAGGAGCGAGCCGCCACGAGAAATGAAGAG 349
 34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
 :|||
 350 AATCCGCCATTGGGAGGGGTGAAGCCAGAGCCCTGGAGGAATATC 399
 51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67
 |||||
 400 AGCGGGGGCGAGTTAGGCGACTTGTCCCTAATTTTCGATGGCCATACC 449
 67 thisArg...LeuGlyGluProGlnAlaArgMetArgGluGluAsnMetG 83
 :||||
 450 TAATAGGCATATTGACCAATGACCGAGA.....GATGATGTAG 490
 83 luArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLys 99
 |||||
 491 AAAGGTTTGTAGGCGAGATGATGGAATCAAGAGAAAGACTAGGGAACAG 540
 100 GlnLeuSerHisSerLeuArgAlaValSerThrAspProHisHisAs 116
 |||||
 541 CAGATGAGCAGCTATATGCGCTTCCAAACTCCTGAACCT.....GA 581
 116 phisHisAspGluPheCysLeuMetPro 125
 :||||
 582 CAACCATATGACTTTTGCTCATACCT 609

seq_name: /SIDS2/ycgdata/geneseq/geneseq/NA2001.DAT:AAF23529

seq_documentation_block:

ID: AAF23529 standard; DNA: 891 BP.

XX AAF23529;

XX 22-MAR-2001 (first entry)

XX Human NADE DNA.

XX Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
 XX neurogenetic disease; NF-kappaB; ds.

OS Homo sapiens.
 XX WO200075278-A2.
 XX 14-DEC-2000.
 XX 07-JUN-2000; 2000WO-US15621.
 XX 07-JUN-1999; 99US-0327750.
 XX (UYCO) UNIV COLUMBIA NEW YORK.
 XX Sato T;
 XX WPI; 2001-061707/07.
 XX New p75-neurotrophin receptor-associated cell death executor (NADE) and
 XX the gene encoding NADE, useful for modulating the activity of p75NTR
 XX and for detecting neurodegenerative diseases -
 XX Disclosure; Fig 1; 134pp; English.
 XX The present invention relates to a purified polypeptide capable of
 XX binding neurotrophin receptor (p75-NTR). The invention is useful for
 XX binding and modulating the activity of p75NTR. The peptide mediates
 XX NGF-induced apoptosis, which plays an important role in neurogenetic
 XX diseases. The peptide of the invention and p75NTR are useful for
 XX inhibiting NF-kappaB activation in a cell or a subject, for inducing
 XX caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase
 XX and fragment nuclear DNA in a cell by co-expression of (I) and p75-NTR.
 XX Sequence 891 BP; 251 A; 182 C; 224 G; 234 T; 0 other;

alignment_scores:

Quality: 201.00 Length: 124
 Ratio: 2.610 Gaps: 4
 Percent Similarity: 62.097 Percent Identity: 43.548

alignment_block:

US-09-327-750D-32 x AAF23529 ..

Align seg 1/1 to: AAF23529 from: 1 to: 891

4 LysGluLysArg...AlaValAsnSerLeuSerMetGluAsnAlaAsnGl 19
 |||||
 279 AAAGAAAAACACACACAGAAAAAATCTCATCATGCGCAATATTCACCA 328
 19 nGluAsnGluLysGluGlnValAlaAsnLysGlyGluProLeuAlaL 36
 |||||
 329 GGAAACCAAGAGATGGAGCACCCCTATCGAGATCGAGAGGAA...GACC 375
 36 euProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArgArg 52
 |||||
 376 GCCCTTTGGGAGAGGTGAAGCCACCAGCCCTGCAGAAATCGACGGGA 425
 53 PheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMetHisAr 69
 :||||
 426 ...CAGGCTCGCGACTTGGCCCTAATTTTCGATGGCCATACCCAATAG 472
 69 gLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArgIleG 86
 :||||
 473 GCAGATCAATGATGGGATGGGATGGAGATGATATGGAATATTCA 522
 86 lyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGlnLeuSer 102
 |||||
 523 TGGAGGATGAGAGAAATCAGAGAAACTTAGGAGCTGCAGTTGAGG 572
 103 HisSerLeuArgAlaValSerThrAsp...ProProHisHisAspHisH 118
 :|||
 573 AATGCTCGCTATCTTATGGGAGGTCTCTAATACCATGACCATCA 622
 118 sAspGluPheCysLeuMetPro 125

|||||
623 TGATGAATTTGCCTTATGCCT 644

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAF21748

seq_documentation_block:

ID AAF21748 standard; DNA; 917 BP.

XX AC AAF21748;

XX DT 27-MAR-2001 (first entry)

XX DE Human breast and ovarian cancer associated antigen gene SEQ ID 135.
XX KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW nontropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease; ds.

XX OS Homo sapiens.

XX WO200055173-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05881.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM;

XX WPI; 2000-611515/58.

XX P-PSDB; AAB58845.

XX New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases -
XX
PS Claim 1: Page 581-582; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC nontropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiant activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and antagonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemias; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.

XX Sequence 917 BP; 228 A; 203 C; 270 G; 211 T; 5 other;

alignment_scores:

Quality: 185.50

Length: 108

Ratio: 2.769

Gaps: 3

Percent Similarity: 62.037 Percent Identity: 44.444
alignment_block:
US-09-327-750D-32 x AAF21748 ..
Align seg 1/1 to: AAF21748 from: 1 to: 917
19 GlnGluAsnGluGluLysGluGlnValAlaAsnLysGluProLeuAl 35
|||||
306 CAGGAAACCAAGAGATGGAGCAGCCTATGCAGAATGGAGAGAA...GA 352
35 aLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArgA 52
|||||
353 CCGCCCTTTGGGAGGAGTGAAGGCCACCCAGCCTGCAGGAATCGACGG 402
52 rgPheProValArgGlnProIleLeuGlnTyrArgTTrpAspIleMethis 68
|||||
403 GA...CAGGCTCGCCGACTTGCCTTAATTTTCGATGGGCCCATACCCAAT 449
|||||
69 ArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluArgI 85
|||||
450 AGCCAGATCAATGATGGGATGGGTGGAGATGGAGATGATGGAATATT 499
85 eGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGlnLeu 102
: |||||
500 CATGGAGGAGATGAGAGAAATCAGAGAAACTTAGGAGCTGCAGTTGA 549
102 erHisSerLeuArgAlaValSerThrAsp...ProProHisHisAspHis 117
: |||||
550 GGAATTCCTCGCTATCTTATGGGAGGCTCTCTAATCACCATGACCAT 599
118 HisAspGluPheCysLeuMetPro 125
|||||
600 CATGATGAATTTGCCTTATGCCT 623

THIS PAGE BLANK (USPTO)

OM of: US-09-327-750D-32 to: GenEmbl.* out_format : pfs
Date: Mar 11, 2002 3:34 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n_model -DEV=xlh
-O=/sgn2.1/USPTO_spool/US09327750/runat_110154_20319/app_query.fasta_1.1472
-DB=genEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FCAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRX=blosum62 -TRANS=human40.cdi
-LIST=45 -DCCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09327750 -CGNI_1_1_8673
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-32

Query length: 125

Database: GenEmbl.*

Database sequences: 1472140

Database length: -341344837

Search time (sec): 4557.230000

score_list:

Sequence	Strd	Orig	zscore	EScore	Len	Documentation
gb_pr:AF183416	+	653.00	1124.31	2.4e-54	642	AF183416 Homo sapiens ovarian
gb_pr:AF220189	-	653.00	1122.20	3.1e-54	828	AF220189 Homo sapiens uncharact
gb_pr:HS198P4	-	653.00	1091.01	1.7e-52	35714	AL008708 Homo sapiens DNA sequen
gb_pr:AF237783	+	639.00	1098.39	6.6e-53	791	AF237783 Homo sapiens brain-exp
gb_pr:AF251053	+	602.50	1035.82	2.0e-49	744	AF251053 Homo sapiens X-linked
gb_pat:AX078272	+	602.50	1034.27	2.5e-49	898	AX078272 Sequence 76 from Paten
gb_pr:HSV870H8	+	602.50	1004.84	1.1e-47	31321	AF078272 Human DNA sequence fr
gb_pr:AL133348	-	602.50	1002.69	1.4e-47	40584	AL133348 Human DNA sequence fr
gb_htg:HSU08081	+	602.50	1002.60	1.4e-47	41029	AL022169 Homo sapiens chromos
gb_pr:AF097439	+	496.00	851.36	3.8e-39	785	AF097439 Mus musculus brain exp
gb_pr:AF097437	+	448.50	760.49	4.4e-34	2269	AF097437 Mus musculus Bex1 pr
gb_pr:AF097438	+	446.50	765.31	2.4e-34	835	AF097438 Mus musculus brain exp
gb_pr:AF051347	+	446.50	765.28	2.4e-34	838	AF051347 Mus musculus brain exp
gb_pr:BC003254	+	384.50	647.86	8.3e-28	2901	BC003254 Mus musculus, Similar
gb_sts:G24641	-	348.00	599.29	4.2e-25	504	G24641 human STS WI-11354, sequ
gb_pr:HSV351F8	-	216.50	334.73	2.3e-10	45678	AF07119 Human DNA sequence fr
gb_pr:AK000959	+	202.50	340.49	1.1e-10	1229	AK000959 Homo sapiens cDNA FL
gb_pat:AX100231	+	202.50	339.63	1.2e-10	1364	AX100231 Sequence 32 from Pat
gb_pr:HS635G19	+	202.50	307.05	7.9e-09	69648	AL035494 Human DNA sequence f
gb_pr:HUMOGC	+	201.00	340.57	1.1e-10	891	M38188 Human unknown protein fr
gb_pr:AF187064	+	201.00	340.57	1.1e-10	891	AF187064 Homo sapiens p75NTR-as
gb_pr:BC003190	+	199.50	338.94	1.3e-10	793	BC003190 Homo sapiens, p75NTR-as
gb_pr:AF187066	+	186.00	316.65	2.3e-09	700	AF187066 Mus musculus p75NTR-as
gb_pr:AF097440	+	186.00	315.00	2.9e-09	854	AF097440 Mus musculus brain exp
gb_sts:G35294	-	170.00	292.18	5.4e-08	477	G35294 human STS SHGC-37409, se
gb_pr:AF187065	+	169.00	289.75	7.3e-08	519	AF187065 Rattus norvegicus p75N
gb_sts:G27208	-	138.50	238.78	5.0e-05	421	G27208 MARC 4953-4954:991939031
gb_sts:G23964	-	103.50	179.33	0.1033	372	G23964 human STS WI-15922, sequ
gb_om:AF023169	+	97.50	147.57	6.07	4923	AF023169 Canis familiaris type
gb_om:MMU05823	+	89.00	130.03	57.60	6945	U05823 Mus musculus pericentr
gb_pr:HSU31906	+	89.00	130.00	57.77	6965	X1906 Homo sapiens golgin-245
gb_pr:HSUGOLGIN	+	89.00	129.19	64.12	7683	X2834 H.sapiens mRNA for golg
gb_pr:HSU41740	+	89.00	129.18	64.23	7695	U41740 Human trans-Golgi p230
gb_ba:AE000565	+	89.00	125.65	100.94	11776	AE000565 Helicobacter pylori
gb_pr:AL136998	+	88.00	100.96	2.4e+03	186218	AL136998 Mouse DNA sequence
gb_pr:AB062945	+	87.00	139.98	16.06	1376	AB062945 Macaca fascicularis b
gb_pl:AF332874	+	86.50	135.20	29.65	2207	AF332874 Oryza sativa phosphol
gb_pr:RATNESTIN	+	86.50	126.99	84.99	5946	M34384 Rat nestin mRNA, comple
gb_pr:AF332860S2	+	86.50	104.48	1.5e+03	89994	AF332861 Mus musculus diadend
gb_ba:LDMOFA	-	86.00	122.32	154.69	9413	Z25774 Leptothrix discophora m

gb_htg:AC068585 + 86.00 105.66 1.3e+03 70358 AC068585 Homo sapiens clon
gb_htg:AC093313 - 86.00 100.00 2.7e+03 139315 AC093313 Trypanosoma cruz
gb_ro:AB041591 - 85.00 134.99 30.49 1657 AB041591 Mus musculus brain
gb_htg:AC013960 + 85.00 102.53 2.0e+03 83321 AC013960 Drosophila melano
gb_in:AC009346 + 85.00 96.70 4.1e+03 169448 AC009346 Drosophila melar

seq_name: gb_pr:AF183416

seq_documentation_block:

LOCUS AF183416 642 bp mRNA PRI 02-SEP-2000
DEFINITION Homo sapiens ovarian granulosa cell 13.0 kDa protein hGR74 homolog
mRNA, complete cds.

ACCESSION AF183416

VERSION AF183416.1 GI:9963770

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 642)

AUTHORS Peng, Y., Qian, B., Tu, Y., Xu, S., Han, Z., Fu, G. and Chen, Z.

TITLE A novel gene expressed in human adrenal gland

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 642)

AUTHORS Peng, Y., GU, W., Huang, C., Xu, S., Han, Z., Fu, G. and Chen, Z.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-1999) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China

FEATURES

source

1..642

/organism="Homo sapiens"

/db_xref="taxon:9606"

/tissue_type="adrenal gland"

3..380

/note="HGR74-h protein"

/codon_start=1

/evidence="not_experimental"

/product="ovarian granulosa cell 13.0 kDa protein hGR74

homolog"

/protein_id="AAG09685.1"

/db_xref="GI:9963771"

/translation="MESKEKRAVSLSMENAOENEEKEQVANKGEPLALPLDAGEYC

VPRGNRRFRVROPILQYRWDMMHRLGEPQARMRENMERIGEEVRLMEKLRLKRLS

HSLRAVSTDPHPHHDHDECLMP"

BASE COUNT 204 a 118 c 157 g 163 t

ORIGIN

alignment_scores:

Quality: 653.00 Length: 125

Ratio: 5.266 Gaps: 0

Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:

US-09-327-750D-32 x AF183416

Align seg 1/1 to: AF183416 from: 1 to: 642

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17

3 ATGCAGTCCAAAGAGAAACGAGCAGTAAACAGTCTCAGCATGGAATGC 52

17 aAsnGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProL 34

53 CAACCAAGAAATGAAGAAAGAGCAGAGTGTCTAATAAAGGGGGCCCT 102

34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50

103 TGGCCCCCCTTGGATGCTGGTGAATCTGTGTGCCTAGAGGAATCGT 152

51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67

||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

153 AGGCGGTTCCGGTTAGGACGCCCATCTCGCAGTATAGATGGGATATGAT 202
67  tHisArgLeuGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
|||||
203 GCATAGGCTTGAGACACACAGGCAAGGATGAGAGAGATATGGAAA 252
84  rGileGlyGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
|||||
253 GGATTGGGAGGAGTGACAGACGTGATGGAAAAGCTGAGGCAAAAGCAG 302
101 LeuSerHisSerLeuArgAlaValSerThrAspProProHisHisAspHi 117
|||||
303 TTGAGTCATAGTCGCGGGCAGTCAGCACTGACCCCTCACCATGACCA 352
117 sHisAspGluPheCysLeuMetPro 125
|||||
353 TCATGATGAGTTTTCCTTATGCCC 377

```

seq_name: gb_pr:AF220189

```

seq_documentation_block:
LOCUS AF220189 828 bp mRNA PRI 04-MAY-2000
DEFINITION Homo sapiens uncharacterized hypothalamus protein HBEX2 mRNA,
complete cds.
ACCESSION AF220189
VERSION AF220189.1 GI:7689028
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Xiao,H., Song,H., Gao,G., Ren,S., Chen,Z. and Han,Z.
TITLE A novel gene expressed in human hypothalamus
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 828)
AUTHORS Xiao,H., Song,H., Gao,G., Ren,S., Chen,Z. and Han,Z.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China
FEATURES
source
1..828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="hypothalamus"
167..544
/codon_start=1
/product="uncharacterized hypothalamus protein HBEX2"
/protein_id="AAF67654.1"
/db_xref="GI:7689029"
/translation="WESKEKRAVNSLSMENAOENKEEOVANKGEPLALPLDAGEYCV
VRGNRRFRFRVQPIQYRWDMHRLGEPQARMRENERMERIGEVRQLMEKLRKQLS
HSLRAVSTDPHHDHDFCLMP"
BASE COUNT 256 a 170 c 221 g 181 t
ORIGIN

```

alignment_scores:
 Quality: 653.00 Length: 125
 Ratio: 5.266 Gaps: 0
 Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block: *
 US-09-327-750D-32 x AF220189 ..

Align seg 1/1 to: AF220189 from: 1 to: 828

```

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
167 ATGAGGTCCAAAGAAACGACGACGATTAACAGTCTCAGCATGGAAATGC 216
17 aAsnGlnGluAsnGluLysGluGlnValAlaAsnLysGluGluProL 34

```

```

217 CAACCAAGAAAATGAAGAAAAGGAGCAAGTTGCTAATAAAGGGAGCCCT 266
34 euAlaLeuProLeuAspAlaGlyClyTyrCysValProArgGlyAsnArg 50
|||||
267 TGGCCCTCCCTTTGGATGCTGGTGAATACTGTGTGCCTAGAGGAAATCGT 316
51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67
|||||
317 AGCGGTTCCGCGTTAGGACGCCCATCTGCAGTATAGATGGGATATGAT 366
67 tHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
|||||
367 GCATAGGCTTGAGAACACACAGGCAAGGATGAGAGAGAGAATATGAAA 416
84 rGileGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
|||||
417 GGATTGGGAGGAGGTGAGACAGCTGATGGAAAAGCTGAGGGAAGACAG 466
101 LeuSerHisSerLeuArgAlaValSerThrAspProProHisHisAspHi 117
|||||
467 TTGAGTCATAGTCGCGGGCAGTCAGCACTGACCCCTCACCATGACCA 516
117 sHisAspGluPheCysLeuMetPro 125
|||||
517 TCATGATGAGTTTTCCTTATGCCC 541

```

seq_name: gb_pr:HS198P4

```

seq_documentation_block:
LOCUS HS198P4 35714 bp DNA PRI 16-DEC-2000
DEFINITION Human DNA sequence from clone RPL-198P4 on chromosome Xq22 Contains
a CpG Island, complete sequence.
ACCESSION AL008708
VERSION AL008708.4 GI:4826445
KEYWORDS HTG; CpG Island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Corby,N.
TITLE Direct Submission
JOURNAL Submitted (21-JAN-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk
COMMENT
On May 13, 1999 this sequence version replaced GI:4582117.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
This sequence was generated from part of bacterial clone contigs of
human chromosome X, constructed by the Sanger Centre Chromosome X
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
from the library RPCI-1 constructed at the Roswell Park Cancer
Institute by the group of Pieter de Jong. For further details see
http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
IMPORTANT: This sequence is not the entire insert of clone

```

Percent Similarity: 91.667 Percent Identity: 84.722

alignment_block:

US-09-327-750D-31 x AAC06180

Align seg 1/1 to: AAC06180 from: 1 to: 457

```
1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
243 ATGGAGTCCAAAGAGAGAGCGGTTAAACAATCTCATCTGCGAATGTT 292
|||||
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||
293 CAACACAGGAAATGATGAAAAAGATGAAAAAGGAGCAAGTTGCTAATAAAG 342
|||||
34 lylGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
343 GGGAGCCCTTGGCCCTTACCTTTGAATGTTAGTGAATACTGTGTGCTAGA 392
|||||
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGln.TyrArgT 67
|||||
393 GGAACCCGTAG.CGGTTCCGCGTTAGGCACACAGTATTAGAGCCACCGCT 441
|||||
67 rpAspIleMethis 71
::: |||
442 GCCCAGTGACACAT 455
```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAC06184

seq_documentation_block:

ID AAC06184 standard; cDNA; 451 BP.

AC AAC06184;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 10259.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

OS Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 1; SEQ ID 10259; 71pp + CD-ROM; English.

PS The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be

CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.

XX SQ Sequence 451 BP; 117 A; 110 C; 131 G; 91 T; 2 other;

alignment_scores:

Quality: 261.00 Length: 69
Ratio: 4.143 Gaps: 1
Percent Similarity: 91.304 Percent Identity: 81.159

alignment_block:

US-09-327-750D-31 x AAC06184

Align seg 1/1 to: AAC06184 from: 1 to: 451

```
4 LysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnValAsnGlnG 20
|||||
246 CAAGAGAGAGCGGTTAAACAATCTCATCTGCGAATGTTCAACACAGA 295
|||||
20 uAsnAspGluLysAspGluLysGluGlnValAlaAsnLysGluProL 37
|||||
296 AAATGATGAAAAAGATGAAAAAGGAGCAAGTTGCTAATAAAGGGAGCCCT 345
|||||
37 euAlaLeuProLeuAsnValSerGluTyrCysValProArgGlyAsnArg 53
|||||
346 TGGCCCTACCTTTGAATGTTAGTGAATACTGTGTGCTAGAGGAACCGT 395
|||||
54 ArgArgPheArgValArgGlnProIleLeuGln.TyrArgTyrAspIle 70
|||||
396 AG.CGGTTCCGCGTTAGGCACACAGTATTAGAGGCACCGCTGCCAGTGA 444
|||||
70 ethis 71
|||
445 CACAT 449
```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAC23529

seq_documentation_block:

ID AAC23529 standard; DNA; 891 BP.

XX AAC23529;

XX 22-MAR-2001 (first entry)

XX Human NADE DNA.

XX Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
KW neurogenetic disease; NF-kappaB; ds.

XX Homo sapiens.

XX WO200075278-A2.

XX 14-DEC-2000.

XX 07-JUN-2000; 2000WO-US15621.

XX 07-JUN-1999; 99US-0327750.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Sato T;

XX WPI; 2001-061707/07.

XX New p75-neurotrophin receptor-associated cell death executor (NADE) and
XX the gene encoding NADE, useful for modulating the activity of p75NTR
XX and for detecting neurodegenerative diseases -
XX Disclosure; Fig 1; 134pp; English.

117 sHIsAspHIsHisAspGluPheCysLeuMetPro 128
 |||||
 514 CCATGACCATCATGATGAGTTTCCTTATGCC 547

seq_name: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT:AAC06183

seq_documentation_block:

ID AAC06183 standard; cDNA: 421 BP.

XX AAC06183;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 10258.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPT; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 10258; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.

XX Sequence 421 BP; 110 A; 103 C; 124 G; 82 T; 2 other;

alignment_scores:

Quality: 284.00 Length: 72
 Ratio: 4.303 Gaps: 1
 Percent Similarity: 91.667 Percent Identity: 84.722

alignment_block:

US-09-327-750D-31 x AAC06183 ..

Align seg 1/1 to: AAC06183 from: 1 to: 421

1 MetGluSerLysGluGluArgAlaLeuAsnLeuValGluAsnVa 17
 |||||
 207 ATGGAGTCCAAAGAGACGACGCGTTAAACAATCTCATCGTGGAATAATGT 256

17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
 |||||
 257 CAACACAGGAATATGATGAAAAAGATCAAAAGGAGCAAGTTGCTAATAAG 306

34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
 |||||
 307 GGGAGGCCCTTGGCCCTTACCTTTGAATGTTAGTGAATACTGTGTGCTAGA 356

51 GlyAsnArgArgPheArgValArgGlnProIleLeuGln.TyrArgT 67
 |||||
 357 GGAACCCCTAG.CGGTTCCGCGTTAGGCACACAGTATTAGAGCACCGCT 405

67 rpAspIleMethHis 71

::: |||

406 GCCCAGTGACACAT 419

seq_name: /SID52/gcgdata/geneseq/geneseq/NA2000.DAT:AAC06180

seq_documentation_block:

ID AAC06180 standard; cDNA: 457 BP.

XX AAC06180;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 10255.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPT; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 10255; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.

XX Sequence 457 BP; 120 A; 110 C; 142 G; 84 T; 1 other;

alignment_scores:

Quality: 284.00 Length: 72
 Ratio: 4.303 Gaps: 1

CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.

XX
 SQ Sequence 692 BP; 184 A; 151 C; 199 G; 146 T; 12 other;

alignment_scores:
 Quality: 626.00 Length: 125
 Ratio: 5.174 Gaps: 0
 Percent Similarity: 96.800 Percent Identity: 93.600

alignment_block:
 US-09-327-750D-31 x AAC10889 ..

Align seg 1/1 to: AAC10889 from: 1 to: 692

4 LysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnValAsnGlnG1 20

246 CAAGAGACGACGGTGAACATCTCATCGTGGAAATGTCACACGAG 295

20 uAsnAspGluLysAspGluGlnValAlaAsnLysGlyGluProL 37

296 AATGATGAAAGATGAAAGGAGCAGTGTGCTAATAAGGGAGCCCT 345

37 euAlaLeuProLeuAsnValSerGluTyrCysValProArgGlyAsnArg 53

346 TGCCCTACCTTTGATGTGTGTGAATCTGTGTGCTAGAGGAAATCGT 395

54 ArgArgPheArgValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 70

396 AGCGGTTCCGCTTAGGAGCCCTCTGCAGTATAGATGAGGATATGAT 445

70 tHisArgLeuGlyGluProGlnAlaArgMetArgGluAsnMetGluA 87

446 GCATAGCTTGGAGAACACAGCAGGAGGATGANAAGAGAAATATGAAA 495

87 rgIleGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 103

496 GGATGGGAGGAGGTGAGACAGCTGATGAAAAGCTGAGGGAAGACAG 545

104 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 120

546 TTGAGTCATAGTCTGGGCGGAGTCACACTGACCCCTCACCATGACCA 595

120 sHisAspGluPheCysLeuMetPro 128

596 TCATGATGAGTTTTCNNWATGCC 620

seq_name: /SID62/9c9data/geneseq/geneseq/NA2001.DAT:AAH45143

seq_documentation_block:

ID AAH45143 standard; cDNA: 792 BP.

XX AC AAH45143;

XX DT 07-SEP-2001 (first entry)

XX DE Human brain expressed x-linked protein, hBex, coding sequence.

XX KW Human; brain expressed x-linked protein; cytostatic; auditory; nontropic;

XX KW hBex; dysembryoplasia; hereditary disease; cancer; tumour; deafness;

XX KW x-chromosome-binding mental retardation; lissencephalous disease; ss.

XX Homo sapiens.

OS WO200140286-A1.

XX PD 07-JUN-2001.

XX PF 27-NOV-2000; 2000WO-CN00502.

XX PR 30-NOV-1999; 95CN-0124179.

XX PA (BIOR-) BIORAD GENE DEV LTD SHANGHAI.

XX PI Mao Y, Xie Y;

XX DR WPI: 2001-397944/42.

XX DR P-PSDB: AAB99224.

XX Isolated human brain-expressed X-linked polypeptide used to diagnose
 PT and treat of dysembryoplasia, hereditary diseases, cancer, tumor,
 PT deafness and X-chromosome-binding mental retardation

XX PS Claim 5: Page 22; 30pp; Chinese.

XX The present sequence is the coding sequence for a human brain-expressed
 CC x-linked protein (hBex). hBex and its coding sequence are useful in the
 CC diagnosis and treatment of dysembryoplasia, hereditary diseases, cancer,
 CC tumours, deafness, x-chromosome-binding mental retardation and
 CC lissencephalous disease. hBex is also useful for screening mimics,
 CC agonists, or inhibitors, and in peptide fingerprinting identification.
 CC hBex coding sequence can be used as primers or probes, or in producing
 CC gene chips or microarrays.

XX SQ Sequence 792 BP; 214 A; 172 C; 219 G; 187 T; 0 other;

alignment_scores:
 Quality: 592.50 Length: 128
 Ratio: 4.857 Gaps: 1
 Percent Similarity: 95.312 Percent Identity: 87.500

alignment_block:

US-09-327-750D-31 x AAH45143 ..

Align seg 1/1 to: AAH45143 from: 1 to: 792

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17

173 ATGAGTCCAAAGAGAAACTAGCAGTAAACAGTCTCAGCATGGAAATGC 222

17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34

223 CAACCAAGAAAAAT.....GAAGAAAGAGCAAGTTGCTAATAAAG 263

34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50

264 GGGAGCCCTTGGCCCTCCCTTTGGATGCTGCTGAATACTGTGTCCTAGA 313

51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67

314 GGAATCGTAGGCGGTTCGCCGTTAGCAGCCCATCTGCAGTATAGATG 363

67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84

364 GGATATGATGATAGGCTTGGAGAACCCAGGCAAGGATGAGAGAGAGA 413

84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100

414 ATATGAAAGCATTTGGGAGGGGTGAGACAGCTGATGGAAAGCTGAGG 463

101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117

464 GAAAGACAGTTGATGATAGTCTCGCGGCGAGTCAGCACTGACCCCTCA 513

```

84  snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 100
|||||
457 ATATGGAAGCATTTGGGAGAGGTGAGACAGCTGATGGAAGCTGAGG 506
|||||
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH1 117
|||||
507 GAAAGCAGTTGAGTCATAGCTCTCGGGCAGTCAGCACTGACCCCTCA 556
|||||
117 SHisAspHisAspGluPheCysLeuMetPro 128
|||||
557 CCATGACCATCATGATGAGTTTTCNNWATGCCCC 590
|||||

```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT: AAC03879

seq_documentation_block:
ID AAC03879 standard; cDNA; 698 BP.

```

XX AAC03879;
AC
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 3877.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
DR P-PSDB; AAG03873.
XX

```

New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures - Claim 1; SEQ ID 3877; 71pp + CD-ROM; English.

The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. An ORF has been identified within the sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.

XX Sequence 698 BP; 187 A; 151 C; 210 G; 139 T; 11 other;

alignment_scores:			
Quality:	649.00	Length:	128
Ratio:	5.234	Gaps:	0
Percent Similarity:	96.875	Percent Identity:	95.312

alignment_block:

```

US-09-327-750D-31 x AAC03879
Align seg 1/1 to: AAC03879 from: 1 to: 698

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
243 ATGAGGTCCAAAGAGGAAACGAGCGTTAAACAATCTCATCGTGGAAATGT 292
|||||
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||
293 CAACCCAGGAAATCATCAAAAGATGAAAAGAGGACCAAGTGCCTAATAAAG 342
|||||
34 lYGlupProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
343 GGGAGCCCTTGGCCCTTACCTTTGATGTYTRGTGAATACTGTGTCCTAGA 392
|||||
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
|||||
393 GGAATCGTAGCGGTTCCGCGTTAGCGACCCCATCTCGCAGTATAGATG 442
|||||
67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
|||||
443 GGATATGATCATAGGCTTGGAGAACCCACAGCAAGGATGANAGAAGAGA 492
|||||
84 snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 100
|||||
493 ATATGGAAGGATTGGGGAGGAGGTGAGACAGCTGATGGAAGAGCTGAGG 542
|||||
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH1 117
|||||
543 GAAAGCAGTTGAGTCATAGCTCTCGGGCAGTCAGCACTGACCCCTCA 592
|||||
117 SHisAspHisAspGluPheCysLeuMetPro 128
|||||
593 CCATGACCATCATGATGAGTTTTCNNWATGCCCC 626
|||||

```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT: AAC10889

seq_documentation_block:

ID AAC10889 standard; cDNA; 692 BP.

```

XX AAC10889;
AC
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 14964.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX

```

New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures - Claim 1; SEQ ID 14964; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from

CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 858 BP; 236 A; 175 C; 241 G; 206 T; 0 other;

alignment_scores:
 Quality: 666.00 Length: 129
 Ratio: 5.203 Gaps: 1
 Percent Similarity: 99.225 Percent Identity: 99.225

alignment_block:
 US-09-327-750D-31 x AAI60367 ..

Align seg 1/1 to: AAI60367 from: 1 to: 858

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
 215 ATGGAGTCCAAAGAGAAAGAGCGGTTAAACAAATCTCATCTGGAATAATGT 264
 17 lAsnGlnGluAsnAspGluLysAspGluGlnValAlaAsnLysG 34
 265 CAACACAGAAATGATGAAAAAGATGAAAGAGCAAGTTGCTAATAAAG 314
 34 lylGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
 315 GGCAGCCCTGGGCCCTACCTTTGAATGTTAGTGAATACTGTGTCCTAGA 364
 51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
 365 GGAACACGTTAGCGGTTCCGCGTTAGCGACCCATCTCGAGTATAGATG 414
 67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu 84
 415 GGACATAATGCATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGGAGA 464
 84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
 465 ATATGGAAGGATTGGGAGGAGGTGAGACAGCTGATGGAAGCTGAGG 514
 101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
 515 GAAAGACAGTTGAGTTCATAGTCTCGCGGCGAGTCAGCACTGATCCCCCTCA 564
 117 sHisAspHisHisAspGluPheCys LeuMetPro 128
 565 CCATGACCATCAGATGAGTTTGGCCCTTATGCC 599

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT: AAC03880

seq_documentation_block:

ID AAC03880 standard; cDNA; 662 BP.
 XX
 AC AAC03880;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein 5' EST, SEQ ID NO: 3878.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

gene therapy; chromosome mapping; ss.

Homo sapiens.

EP1033401-A2.

06-SEP-2000.

21-FEB-2000; 2000EP-0200610.

26-FEB-1999; 99US-0122487.

(GEST) GENSET.

Dumas Milne Edwards J, Duclert A, Giordano J;

WPI; 2000-500381/45.

P-PSDB; AAG03874.

New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 diagnostic, forensic, gene therapy and chromosome mapping procedures -
 Claim 1; SEQ ID 3878; 71pp + CD-ROM; English.

The present sequence is one of a large number of 5' ESTs derived from
 mRNAs encoding secreted proteins. An ORF has been identified within the
 sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 derived from 30 different tissues. EST sequences usually correspond
 mainly to the 3' untranslated region (UTR) of the mRNA because they are
 often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 well suited for isolating cDNA sequences derived from the 5' ends of
 mRNAs and even in those cases where longer cDNA sequences have been
 obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 mRNAs with intact 5' ends and can therefore be used to obtain full length
 cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 gene therapy and chromosome mapping procedures. They are used to obtain
 upstream regulatory sequences and to design expression and secretion
 vectors.

Sequence 662 BP; 177 A; 144 C; 192 G; 137 T; 12 other;

alignment_scores:

Quality: 649.00 Length: 128
 Ratio: 5.234 Gaps: 0
 Percent Similarity: 96.875 Percent Identity: 95.312

alignment_block:

US-09-327-750D-31 x AAC03880 ..

Align seg 1/1 to: AAC03880 from: 1 to: 662

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
 207 ATGGAGTCCAAAGAGAAAGAGCGGTTAAACAAATCTCATCTGGAATAATGT 256
 17 lAsnGlnGluAsnAspGluLysAspGluGlnValAlaAsnLysG 34
 257 CAACACAGAAATGATGAAAAAGATGAAAGAGCAAGTGCCTAATAAAG 306
 34 lylGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
 307 GGCAGCCCTGGGCCCTACCTTTGATGCTGCTGAATACTGTGTCCTAGA 356
 51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
 357 GGAATCGTAGGCGGTTCCGCGTTAGCGACCCCATCTCGAGTATAGATG 406
 67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu 84
 407 GGATATGATGATAGGCTTGGAGAACCCACAGGCAAGGATGANAGAAGAGA 456

KW menstrual cycle disorder; bacterial infection; ss.
XX Homo sapiens.
OS
XX WO200107471-A2.
PN
XX
XX 01-FEB-2001.
PD
XX
XX 21-JUL-2000; 2000WO-US19948.
PF
XX
XX 21-JUL-1999; 99US-0145075.
PR
XX 08-SEP-1999; 99US-0153129.
PR
XX 10-NOV-1999; 99US-0164647.
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
PI Azimzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;
PI
XX
XX WPI: 2001-112727/12.
DR
XX P-PSDB: AAB60474.
XX
XX Human cell cycle and proliferation proteins and polynucleotides are
PT used to treat, diagnose and prevent immune, developmental and cell
PT signaling disorders and cell proliferative disorders including cancer -
PT
XX Claim 5; Page 181-182; 205pp; English.
XX
XX Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human
CC cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.
CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
CC associated with decreased expression of functional CCYPR, while CCYPR
CC antagonists are used to treat diseases or conditions associated with
CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
CC that specifically bind to CCYPR, and in drug screening methods to
CC identify compounds that modulate the activity of CCYPR. CCYPR
CC nucleotides can be used to generate transgenic animal models of human
CC disease, and can be used in gene therapy in target cells with genetic
CC abnormalities with respect to the expression of CCYPR for the
CC treatment or prevention of a disorder associated with CCYPR.
CC Diseases which can be diagnosed, treated and prevented using CCYPR
CC proteins, nucleic acids, agonists or antagonists include immune,
CC developmental and cell signalling disorders, and cell proliferative
CC disorders including cancer. Specific examples of these disorders
CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
CC diabetes mellitus, disorders of the menstrual cycle and infections
CC caused by bacteria.
XX
SQ Sequence 898 BP; 250 A; 186 C; 251 G; 211 T; 0 other;

alignment_scores:
Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-327-750D-31 x AAF59611 ..
Align seg 1/1 to: AAF59611 from: 1 to: 898
1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuValGluAsnVa 17
254 ATGGAGTCCAAAGAGGAGGAGCGGTAAACAATCTCATCGTGGAATGT 303
17 lasnGlnGluAsnAspGluLysAspGluGlnValAlaAsnLysG 34
304 CAACGAGGAAATGATGAAAAAGATGAAAGGAGCAAGTTGCTATAAAG 353
34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50

354 GGGAGCCCTTGGCCCTACCTTTGAATGTTAGTGAATAC'TGTGTGCCTAGA 403
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
404 GGAACCGTAGCGGTTCGGCGTTAGGCAGCCCATCTGCAGTATAGATG 453
67 pAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluGluA 84
454 GGACATAATGCATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGGAGA 503
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
504 ATATGAAAGGATTGGGAGGAGGTGAGACACCTGATGAAAAGCTGAGC 553
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 117
554 GAAAGCAGTTGAGTCATAGTCTGCGGCGAGTCAGCAC'TGATCCCCCTCA 603
117 shHisAspHisAspGluPheCysLeuMetPro 128
604 CCATGACCATCACGATGAGTTTGGCCTTATGCC 637
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAI60367
seq_documentation_block:
ID AAI60367 standard; cDNA; 858 BP.
XX
XX AAI60367;
AC
XX 22-Oct-2001 (first entry)
DT
XX Human polynucleotide SEQ ID NO 4356.
DE
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-prager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
XX Homo sapiens.
OS
XX WO200153312-A1.
PN
XX 26-JUL-2001.
PD
XX 26-DEC-2000; 2000WO-US34263.
PF
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-052317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
PI
XX WPI: 2001-442253/47.
DR P-PSDB: AAM41211.
DR
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PT
XX Claim 1; SEQ ID NO 4356; 10078pp; English.
PS
XX The invention relates to human nucleic acids (AAI57798-AAI61369), and
XX CC

251 GGAACCGTAGCGGTTCCGCGTTAGCAGCCCATCTCGCAGTATAGATG 300
67 pAspIleMetHisArgLeuGluProGlnAlaArgMetArgGluGluA 84
|||||
301 GGACATATGCATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGAGAGA 350
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
|||||
351 ATATGGAAGGATTGGGGAGAGGTGACAGCTGATGGAAAGCTGAGG 400
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||
401 GAAAGCAGTTGAGTCATAGTCTCGCGGCAGTCAGCAGCTGATCCCCCTCA 450
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
451 CCATGACCATCAGCATGAGTTTGCCTTATGCCC 484

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAI58581

seq_documentation_block:

ID AAI58581 standard; cdna; 862 BP.

XX

AC AAI58581;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 784.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.

XX

OS Homo sapiens.

XX

PN WO200153312-A1.

XX

PD 26-JUL-2001.

XX

PF 26-DEC-2000; 2000WO-US34263.

XX

PR 21-JAN-2000; 2000US-0488725.

PR

PR 25-APR-2000; 2000US-0552317.

PR

PR 09-JUL-2000; 2000US-0598042.

PR

PR 19-JUL-2000; 2000US-0620312.

PR

PR 03-AUG-2000; 2000US-0653450.

PR

PR 14-SEP-2000; 2000US-0662191.

PR

PR 19-OCT-2000; 2000US-0693036.

PR

PR 29-NOV-2000; 2000US-0727344.

XX

XX (HYSE-) HYSEQ INC.

PA

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX

DR WPI; 2001-44253/47.

DR P-PSDB; AAM39425.

XX

XX Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

XX

PS Claim 1; SEQ ID NO 784; 10078pp; English.

XX

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and

CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX

SQ Sequence 862 BP; 235 A; 183 C; 238 G; 206 T; 0 other;

alignment_scores:

Quality: 679.00 Length: 128

Ratio: 5.305 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-31 x AAI58581

Align seg 1/1 to: AAI58581 from: 1 to: 862

1 MetGluSerLysGluArgAlaLeuAsnLeuIleValGluAsnVa 17

|||||
226 ATGGAGTCCAAAGAGGAACGAGCGTTAAACAATCTCATCTCGTGGAAATGT 275

17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34

|||||
276 CAACGAGGAAATGATGAAAAGATGAAAAGGAGCAAGTTGCTAATAAAG 325

34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50

|||||
326 GGGAGCCCTTGGCCCTACCTTTGAATGTTAGTGAATACTGTGTGCCCTAGA 375

51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67

|||||
376 GGAACCGTAGCGGTTCCGCGTTAGGAGCCCATCTCGCAGTATAGATG 425

67 pAspIleMetHisArgLeuGluProGlnAlaArgMetArgGluGluA 84

|||||
426 GGACATAATGCATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGGAGA 475

84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100

|||||
476 ATATGGAAGGATTGGGAGAGGAGTGAGACAGCTGATGGAAAAGCTGAGG 525

101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117

|||||
526 GAAAAGCAGTTGAGTCATAGTTTGGGGCAGTCAGCAGCTGATCCCCCTCA 575

117 sHisAspHisHisAspGluPheCysLeuMetPro 128

|||||
576 CCATGACCATCAGCATGAGTTTGCCTTATGCCC 609

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF59611

seq_documentation_block:

ID AAF59611 standard; cdna; 898 BP.

XX

AC AAF59611;

XX

DT 24-APR-2001 (first entry)

XX

DE Human cell cycle and proliferation protein CCYPR-22 cdna, SEQ ID NO:76.

XX

XX Cell cycle and proliferation protein; CCYPR; human; agonist;

KW antagonist; gene therapy; detection; gene therapy;

KW transgenic animal disease model; immune disorder;

KW developmental disorder; cell signalling disorder;

KW cell proliferative disorder; cancer; tumour; anaemia;

KW arteriosclerosis; asthma; allergy; diabetes mellitus;


```
31 aAsnLysGlyLeuProLeuAlaLeuProLeuAsnValSerGluTyrCysv 48
   ::::::::::::::::::::|
883 ACAGCGTGGAGACCACTCGAGCAACCACTGGAGCTACACCAAC... 928
   ::::::::::::::::::::|
48 alProArgGlyAsnArgArgPheArgValArgGlnProIleLeuGln 64
   ::::::::::::::::::::|
929 .....CGAGGGCCAGCGTGGAGCGCGCATCGTCCCAAGGCGCATCAG 973
   ::::::::::::::::::::|
65 TyrArgTirpAsp..... 68
   ::::::::::::::::::::|
974 CTTCCGCGTGGAGCTGAACCTACCAAGCAGCAGCGTGGCCAGAGT 1023
   ::::::::::::::::::::|
69 .....IleMetHisArgLeuGlyGluPro...GlnAlaArgMetArgGluG 83
   ::::::::::::::::::::|
1024 GGGGCACCAAGCAGCGGCAACAGCAGCTTCAACACCGCCAGCGC... 1069
   ::::::::::::::::::::|
83 luAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99
   ::::::::::::::::::::|
1070 .....CGGTACCTGTAACGCGCAACGTCGCTACAAACAACGTGGG 1108
   ::::::::::::::::::::|
100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValSe 112
   ::::::::::::::::::::|
1109 CACCGCGGCATCTAGCAGCTGAAGCCACCAACAGCTTCGTGCT..... 1153
   ::::::::::::::::::::|
112 rThrAspProProHisHisAspHis 121
   ::::::::::::::::::::|
1154 ....GAACAACGACACCATCGCCACCAT 1177
   ::::::::::::::::::::|

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-471-033-17

seq_documentation_block:
; Sequence 17, Application US/08471033
; Patent No. 5770696
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,033
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: P-40,403
```

```
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV7 - SOLv3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2655 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..2652
; OTHER INFORMATION: /note= "Maize optimized DNA
; OTHER INFORMATION: sequence for 100 kd VIPIA(a) protein from AB78"
US-08-471-033-17
```

```
alignment_scores:
  Quality: 87.00      Length: 126
  Ratio: 1.338      Gaps: 6
  Percent Similarity: 51.587      Percent Identity: 27.778

alignment_block:
US-09-327-750D-31 x US-08-471-033-17 ..

Align seg 1/1 to: US-08-471-033-17 from: 1 to: 2655

15 GluAsnValAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAl 31
   ::::::::::::::::::::|
921 GAACGTGAGCATGGAGAAGTGATCTGAGCCCCCAGCAGCACTGAGCA 970
   ::::::::::::::::::::|
31 aAsnLysGlyLeuProLeuAlaLeuProLeuAsnValSerGluTyrCysv 48
   ::::::::::::::::::::|
971 ACAGCGTGGAGAGCCACTCGAGCAGCAACCTGGAGCTACACCAAC... 1016
   ::::::::::::::::::::|
48 alProArgGlyAsnArgArgPheArgValArgGlnProIleLeuGln 64
   ::::::::::::::::::::|
1017 .....CGAGGGCCAGCGTGGAGCGCGCATCGTCCCAAGGCGCATCAG 1061
   ::::::::::::::::::::|
65 TyrArgTirpAsp..... 68
   ::::::::::::::::::::|
1062 CTTCCGCGTGGAGCTGAACCTACCAAGCAGCAGCGTGGCCAGAGT 1111
   ::::::::::::::::::::|
69 .....IleMetHisArgLeuGlyGluPro...GlnAlaArgMetArgGluG 83
   ::::::::::::::::::::|
1112 GGGGCACCAAGCAGCGGCAACAGCAGCTTCAACACCGCCAGCGC... 1157
   ::::::::::::::::::::|
83 luAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99
   ::::::::::::::::::::|
1158 .....CGGTACCTGTAACGCGCAACGTCGCTACAAACAACGTGGG 1196
   ::::::::::::::::::::|
100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValSe 112
   ::::::::::::::::::::|
1197 CACCGCGGCATCTAGCAGCTGAAGCCACCAACAGCTTCGTGCT..... 1241
   ::::::::::::::::::::|
112 rThrAspProProHisHisAspHis 121
   ::::::::::::::::::::|
1242 ....GAACAACGACACCATCGCCACCAT 1265
   ::::::::::::::::::::|
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 9..2564
; OTHER INFORMATION: /note= "Maize optimized sequence
; OTHER INFORMATION: encoding VIPIA(a) with the Bacillus secretion signal removed
; OTHER INFORMATION: contained in pCIB5526"
US-08-469-334-35

alignment_scores:
    Quality: 87.00      Length: 126
    Ratio: 1.338       Gaps: 6
    Percent Similarity: 51.587      Percent Identity: 27.778

alignment_block:
US-09-327-750D-31 x US-08-469-334-35 ..
Align seg 1/1 to: US-08-469-334-35 from: 1 to: 2576

15 GluAsnValAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAl 31
833 GAACGTGAGCATGGAGAGGTGATCTGTGAGCCCAACGAGACCTGAGCA 882

31 aAsnLysGlyGluProLeuAlaLeuProLeuAsnValSerGluTyCysV 48
883 ACAGCGTGGAGACCCACTCGAGCACCACACTGGAGGTACACCAACAC... 928

48 alProArgGlyAsnArgArgPheArgValArgGlnProIleLeuGln 64
929 .....CGAGGCGCCAGCGTGGAGCGCGGCATCGGTCCCAAGGGCATCAG 973

65 TyrArgTrpAsp..... 68
974 CTTCCGCGTGGAGCGTGAACCTACCAGCAGCAGCGACCGTGGCCAGGAGT 1023

69 .....IleMetHisArgLeuGlyGluPro...GlnAlaArgMetArgGluG 83
1024 GGGGCACACGACCGGACACACACCGCGAGTTCACACCGCCAGCGC.... 1069

83 luAsnMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeu 99
1070 .....CGGCTACCTGACGCCCAACGTCGCTACACCAACGCTGGG 1108

100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValSe 112
1109 CACCGCGGCATCTAGCAGCTGAAGCCACACCGACCGACCGTTCGTGCT.... 1153

112 rThrAspProProHisHisHisHisHis 121
1154 ....GAACAACGACACCATCGCCACCAT 1177

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-300-529-35

seq_documentation_block:
; Sequence 35, Application US/09300529
; Patent No. 6066783
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
```

```
alignment_scores:
    Quality: 87.00      Length: 126
    Ratio: 1.338       Gaps: 6
    Percent Similarity: 51.587      Percent Identity: 27.778

alignment_block:
US-09-327-750D-31 x US-09-300-529-35 ..
Align seg 1/1 to: US-09-300-529-35 from: 1 to: 2576

15 GluAsnValAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAl 31
833 GAACGTGAGCATGGAGAGGTGATCTGTGAGCCCAACGAGACCTGAGCA 882
```

/note= "Maize optimized sequence
encoding VIPIA(a) with the Bacillus secretion signal removed
contained in pCIB5526"

65 TyrArgTirPasp..... 68
 ||| :::
 974 CTTGGCGGTGAGCGTGAACCTACCAGCACAGGACCGTGGCCCGAGGT 1023
 69 TlometHisArnLeuGluGluPro GlnAlaArnMetArgGluG 83

69 ...llemethisargLeuGlyIuPro...GlnAlaArgMetArgGluG 83
 :|||:||||| :|||: |||: |||: |||:
 1024 GGGGCACCAGCACCGGCACACGCCAGTTCACACCGCCAGC... 1069
 83 LuAsnMetGluArgThrGlyGluValArgGlnLeuMetGluLysLeu 99

1070CGGCTACCTGAACGCCAACGTGCGCTACAAACAACGTGGG 1108
100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValSer 112
: :::::|||||
1109 CACGGGGCCATCTACGACGCTGAAGCCCAACCAACAGCTTCGTGCT..... 1153

1109 CACGGCGGCATCTACGACGCTGAGCCACACACAGCTTCGTGCT..... 1153

```

1154 ... GAACACGACCATCGCCACCAT 1177
      :::      |||||      |||||
name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-469-334-35

```

```
name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-469-334-35
documentation_block:
sequence 35, Application US/08469334
tent No. 5990383
```

documentation_block:
quence 35, Application us/08469334
tent No. 5990383
GENERAL INFORMATION:
APPLICANT: Warren, Gregory W
APPLICANT: Kozel, Michael C

GENERAL INFORMATION:

APPLICANT: Warren, Gregory W

APPLICANT: Koziel, Michael G

APPLICANT: Mullins, Martha A

APPLICANT: Nye, Gordon J

APPLICANT: ROZIER, Michael G
APPLICANT: Mullins, Martha A
APPLICANT: Nye, Gordon J
APPLICANT: Carr, Brian
APPLICANT: Desai, Nalini M
APPLICANT: Kostichka, N. Kristy

APPLICANT: Carr, Brian
APPLICANT: Desai, Nalini M
APPLICANT: Kostichka, N. Kristy
APPLICANT: Duck, Nicholas B
APPLICANT: Estruch, Juan J

APPLICANT: Duck, Nicholas B
APPLICANT: Estruch, Juan J
TITLE OF INVENTION: No. 5990383el Pesticidal Proteins and
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:

TITLE OF INVENTION: NO. 5950383el Pesticidal Proteins and
 NUMBER OF SEQUENCES: 50
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CIBA-GEIGY Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne

ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA

STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOURCE: Paterson, Balcer, #10
Version: #1 30B

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,334

SOFTWARE: FACELINK RELEASE #1.0, VERSION #1.00B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,334
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:

FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/463,483
FILING DATE: NUMBETS: SIC 09/2314 504

APPLICATION NUMBER: 08/463,483
FILING DATE:
APPLICATION NUMBER: US 08/314,594
FILING DATE: 09-SEP-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/314,594
FILING DATE: 09-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/218,018
FILING DATE: 23-MAR-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/218,018
FILING DATE: 23-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/037,057
FILING DATE: 25-MAR-1993

APPLICATION NUMBER: US 08/037,057
FILING DATE: 25-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Spruill, W. Murray
REGISTRATION NUMBER: 32,943

NAME: Spruill, W. Murray
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: CGC 1695/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 010-541-86315

REFERENCE/DOCKET NUMBER: CGC 1695/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8615
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 35:

TELEPHONE: 919-541-8613
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 35:

```
1024 GGGCACCAGCAGCGCAACACAGCAGTTCACACCGCCAGCGC.... 1069
      :::::||||| ::::: ||| ||:::|
83 luasnMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeu 99
      :::::||||| ::::: ||| ||:::|
1070 .....CGGCTACCTGACCGCAACGTCGCTACACCAACGTTGG 1108
      :::::||||| ::::: ||| ||:::|
100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValSe 112
      :::::||||| ::::: ||| ||:::|
1109 CACCGGCGCCATCTACGACGTGAAGCCACACAGCTTCGCTGCT.... 1153
      :::::||||| ::::: ||| ||:::|
112 rThrAspProHisHisAspHisHis 121
      :::::||||| ::::: ||| ||:::|
1154 ....GAACAACGACACCATGCCACCAT 1177
```

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-046A-35

```
seq_documentation_block:
; Sequence 35, Application US/08471046A
; Patent No. 5866326
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: Method For Isolating Vegetative Insecticidal
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5866326artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/471,046A
; FILING DATE: 05-JUN-1995
; APPLICATION NUMBER: US 08/463,483
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1695/CIP3/DIV8 - SOLv4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 9..2564
; OTHER INFORMATION: /note= "Maize optimized sequence
; OTHER INFORMATION: encoding VIP1A(a) with the Bacillus secretion signal remov
; OTHER INFORMATION: contained in pCIB5526"
US-08-471-046A-35

alignment_scores:
      Quality: 87.00      Length: 126
      Ratio: 1.338      Gaps: 6
      Percent Similarity: 51.587      Percent Identity: 27.778

alignment_block:
US-09-327-750D-31 x US-08-471-046A-35 ..

Align seg 1/1 to: US-08-471-046A-35 from: 1 to: 2576

15 GluAsnValAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAl 31
   |||::: ::::: |||:::| ::::: |||
833 GAACGTGACATCGAGAGAGGTGATCTGAGCCCAACGACGACCTGAGCA 882
   |||::: ::::: |||:::| ::::: |||
31 aAsnLysGlyProLeuAlaLeuProLeuAsnValSerGluTyrCysv 48
   :::::||||| ::::: ||| ||:::|
883 ACAGCGTGGAGAGCCACTCGAGCACCAACTGGAGCTACACCAACAC... 928
   :::::||||| ::::: ||| ||:::|
48 alProArgGlyAsnArgArgPheArgValArgGlnProIleLeuGln 64
   |||:::||||| ::::: ||| ||:::|
929 .....CGAGGGCGCCAGCGTGGAGCGCGCTGCCCAAGGGCATCAG 973
   |||:::||||| ::::: ||| ||:::|
65 TyrArgTyrAsp.....
   |||:::||||| ::::: ||| ||:::|
974 CTTCCGGCTGAGCGTGAACCTACACGACGACGACCGTGGCCCGAGGAGT 1023
   |||:::||||| ::::: ||| ||:::|
69 .....IleMetHisArgLeuGlyGluPro...GlnAlaArgMetArgGluG 83
   :::::||||| ::::: ||| ||:::|
1024 GGGGCACACGACCGCGCAACACACGACGAGTTCAACACCGCCAGCGC... 1069
   :::::||||| ::::: ||| ||:::|
83 luasnMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeu 99
   |||:::||||| ::::: ||| ||:::|
1070 .....CGGCTACCTGACCGCAACGTCGCTACACCAACGTTGGG 1108
   |||:::||||| ::::: ||| ||:::|
100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValSe 112
   :::::||||| ::::: ||| ||:::|
1109 CACCGGCGCCATCTACGACGTGAAGCCACACAGCTTCGCTGCT.... 1153
   :::::||||| ::::: ||| ||:::|
112 rThrAspProHisHisAspHisHis 121
   :::::||||| ::::: ||| ||:::|
1154 ....GAACAACGACACCATGCCACCAT 1177
```

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-046A-35

```
seq_documentation_block:
; Sequence 35, Application US/08470566B
; Patent No. 5872212
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: NO. 5872212el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
```

```
,
, HYPOTHETICAL: NO
, FEATURE:
, NAME/KEY: CDS
, LOCATION: 9..2564
, OTHER INFORMATION: /note= "Maize optimized sequence
, OTHER INFORMATION: encoding VIPL(a) with the Bacillus secretion signal removed
, OTHER INFORMATION: contained in pCIB526"
US-08-471-044--35
```

```

, STATE: NY
, COUNTRY: USA
, ZIP: 10532
, COMPUTER TYPE: FLOPPY DISK
, MEDIUM TYPE: FLOPPY DISK
, COMPUTER: IBM PC COMPATIBLE
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: PatentIn Release #1.0, Version #1.30B
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/463,483A
,

```

alignment_scores:		
Quality:	87.00	Length: 126
Ratio:	1.338	Gaps: 6
Percent Similarity:	51.587	Percent Identity: 27.778

```
alignment_block:
US-09-327-750D-31 x US-08-471-044-35
```

Align seg 1/1 to: US-08-471-044-35 from: 1 to: 2576

15 GluAsnValAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAl 31
|||:: :: |||::||| :: :: :::::| | |
833 GAACGTGAGCATGGAGAAGGTGATCCTGAGCCCCCAACGAGAACCTTGAGCA 882

31 aaAsnLysGlyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysv 48
.....| | | | |
.....| | | | |

883 ACAGCGTGAGAGCCACTCGAGCACCAACTGGAGCTACACCAACAC.... 928

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

48 **ALP**QWGGGLYASNARGARGARGpneargvalargGlnProIleLeuGln 64

929CGAGGGCCAGCGTGGAGGCCGCATCGGTCCCAAGGCATCAG 973

65 TyrArgTrpAsn 60

..... :: ||

974 CTTGGCGTGAGCGTGAACTACCAACAGCAGCGAGACCGTGCGCCAGGAGT 1023

69IleMetHisArgIleuGlyGluPro...GlnAlaArgMetArgGluG 83

Category	Count	Percentage
1. No response	1	1.0%
2. Very low	1	1.0%
3. Low	1	1.0%
4. Moderate	1	1.0%
5. High	1	1.0%
6. Very high	1	1.0%
7. Extreme	1	1.0%
8. Not applicable	1	1.0%
9. Other	1	1.0%
10. Total	10	100.0%

1024 GGGGCACCAGCACC GCGGCAACACCCAGCCAGTTC AACACCGCCAGCGC... 1069

83 luAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99

[illegible]

1U/0CGGCTACCTGAACGCCAACGTCGGCTACAACAACGTGGG 1108

100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValSe 112

[illegible]

TTT TCC CCG GGG AAA CCC ACC GGC CGG GAG GGT GAA GTT CCT

112 rThrAspProHisHisAspHis 121

1154 GAACAACGACACCATCGCCACCAT 1177

1. The first step is to identify the problem or question that needs to be answered. This involves understanding the context and the specific requirements of the task.

seq_documentation_block:

; Sequence 35, Application US/08463483A

Patent No. 5849870
GENERAL INFORMATICS

APPLICANT: Warren

APPLICANT: KOZIEL, MICHAEL G

APPLICANT: Mullins, Martha A

APPLICANT: Nye, Gordon J
APPLICANT: Carr, Brian

APPLICANT: Desai, Nalini M

; APPLICANT: Kostichka, N. Kris

```

; APPLICANT: DUCK, NICHOLAS B
:
: APPLICANT: Estruch, Juan J

```

TITLE OF INVENTION: NO. 58498

; NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY CORP.

STREET: 7 Skyline Drive

; CITY: Hawthorne

```

;
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,033
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: P-40,403
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV7 - SOLv3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 9..2564
; OTHER INFORMATION: /note= "Maize optimized sequence
; OTHER INFORMATION: encoding VIP1A(a) with the Bacillus secretion signal removed
;
; US-08-471-033-35
;
;
; alignment_scores:
;      Quality: 87.00      Length: 126
;      Ratio: 1.338      Gaps: 6
; Percent Similarity: 51.587      Percent Identity: 27.778
;
; alignment_block:
; US-09-327-750D-31 x US-08-471-033-35
;
; Align seg 1/1 to: US-08-471-033-35 from: 1 to: 2576
;
; 15 GluAsnValAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAl 31
;      |||::: :::: |||::: |||::: |||::: |||::: |||::: |||
; 833 GAACGTGAGCATGGAGAGGTGTCTGAGCCCAACGAGACCTGAGCA 882
;
; 31 aasnLysGlyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysV 48
;      ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; 883 ACAGCGTGAGAGCCACTCGAGCACCACCACTGGAGCTACACCAACAC... 928
;
; 48 alProArgGlyAsnArgArgPheArgValArgGlnProIleLeuGln 64
;      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; 929 .....CGAGGCCCAAGCGGTGGAGCCGCGCATCGTCCCAAGGCGCATCAG 973
;
; 65 TyrArgTrpAsp..... 68
;      ||| :::
; 974 CTTCCGCGTGAGCGTGAACCTACACGACGAGACCGTGGCCCGAGGAGT 1023
;
; 69 .....IleMetHisArgLeuGlyGluPro...GlnAlaArgMetArgGluG 83
;      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
; 1024 GGGGACCAAGCAGCCGCGCAACACCAAGCCAGTTCACACCGCCGCGCG... 1069

```

```

83 luAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99
;      |||::: |||::: |||::: |||::: |||::: |||::: |||
1070 .....CGGTACTGTAACGCAAGCTGCGCTACACCAACGCTGGG 1108
;
100 ArgGluGlnLeu.....SerHisSerLeuArgAlaValSe 112
;      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
1109 CACCGCGGCATCTACGAGCTGAAGCCACCAACGCTTCGTGCT.... 1153
;
112 rThrAspProHisHisHisHis 121
;      ::: ||| ||| ||| ||| ||| ||| ||| ||| |||
1154 ....GAACAACGACACCATCGCCACCAT 1177
;
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-044-35
;
seq_documentation_block:
; Sequence 35, Application US/08471044
; Patent No. 5840868
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mallins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: NO. 5840868el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,044
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,483
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV6 - SOLv3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"

```

```
100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValse 112
      ::::::::::::::|
      ::::::::::::::|
546 CACCGCGGCATCTACGAGTGAAGCCACACAGCTTCGTGCT..... 590
      ::::::::::::::|
112 rThrAspProHisHisAspHisHis 121
      ::::::::::::::|
591 ....GAACAACGACACCATCGCCACCAT 614
```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-300-529-18

```
seq_documentation_block:
; Sequence 18, Application US/09300529
; Patent No. 6066783
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: Genes Encoding Insecticidal Proteins
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6066783artis Corporation
; STREET: 3054 Cordwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/300,529
; FILING DATE: TBA
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/469,334
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/463,483
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-19506L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2004 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

```
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2004
; OTHER INFORMATION: /note= "Maize optimized DNA
; OTHER INFORMATION: sequence for VIP1A(a) 80 kd protein from AB78"
US-09-300-529-18
```

```
alignment_scores:
  Quality: 87.00      Length: 126
  Ratio: 1.338       Gaps: 6
  Percent Similarity: 51.587      Percent Identity: 27.778

alignment_block:
US-09-327-750D-31 x US-09-300-529-18

Align seg 1/1 to: US-09-300-529-18 from: 1 to: 2004

15 GluAsnValAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAl 31
   |||::: :::: |||::: |||::: |||::: |||::: |||::: |||
270 GAACGTGACCATGGAGAAGGTGATCTGTGAGCCGCCAACGAGAACCTGAGCA 319
   ::::::::::::::|
31 aasnLysGlyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysV 48
   ::::::::::::::|
320 ACAGCGTGGAGAGCCACTCGAGCACCAACTGGAGCTACACCAACAC.... 365
   ::::::::::::::|
48 alProArgGlyAsnArgArgArgPheArgValArgGlnProIleLeuGln 64
   |||::: |||::: |||::: |||::: |||::: |||::: |||
366 .....CGAGGCGCCAGCGTGGAGCGCGGCGATCGGTCCCAAGGGCATCAG 410
   |||::: |||::: |||::: |||::: |||::: |||::: |||
65 TyrArgTrpAsp..... 68
   |||::: |||::: |||::: |||::: |||::: |||::: |||
411 CTTCGCGGTGAGCGTGAACCTACCAACGACGACGAGCGTGGCCAGGAGT 460
   |||::: |||::: |||::: |||::: |||::: |||::: |||
69 .....IleMetHisArgLeuGlyGluPro...GlnAlaArgMetArgGluG 83
   ::::::::::::::|
461 GGGGCGCACGACCCGCCACACACAGCCAGGTCAACACCGCCAGCGC.... 506
   ::::::::::::::|
83 luAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99
   |||::: |||::: |||::: |||::: |||::: |||::: |||
507 .....CGGCTACCTGAACGCGCAACGCTCGCTTACCAACACGTTGGG 545
   |||::: |||::: |||::: |||::: |||::: |||::: |||
100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValse 112
   ::::::::::::::|
546 CACCGCGGCATCTACGAGTGAAGCCACACAGCTTCGTGCT..... 590
   ::::::::::::::|
112 rThrAspProHisHisAspHisHis 121
   ::::::::::::::|
591 ....GAACAACGACACCATCGCCACCAT 614
   ::::::::::::::|
```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-471-033-35

```
seq_documentation_block:
; Sequence 35, Application US/08471033
; Patent No. 5770696
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
```

US-08-470-566B-18

alignment_scores:
Quality: 87.00 Length: 126
Ratio: 1.338 Gaps: 6
Percent Similarity: 51.587 Percent Identity: 27.778

alignment_block:

US-09-327-750D-31 x US-08-470-566B-18 ..

Align seg 1/1 to: US-08-470-566B-18 from: 1 to: 2004

15 GluAsnValAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAl 31

270 GAACGTGAGCATGGAGAGGTGATCTGTAGCCCAACGAGAACCTGAGCA 319

31 aAsnLysGlyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysv 48

320 ACAGCGTGGAGAGCCACTCGAGCACCACCACTGGAGGTACACCAACAC. 365

48 alProArgGlyAsnArgArgArgPheArgValArgGlnProIleLeuGln 64

366CGAGGGCGCCAGCGTGGAGCGCGCATCGGTCCCAAGGGCATCAG 410

65 TyrArgTIPAsp..... 68

411 CTTCCGCGTGAGCGTCACTACCAGCAGCAGCGAGCCGTGGCCCGCAGGAT 460

69IleMetHisArgLeuGlyGluPro...GlnAlaArgMetArgGluG 83

461 GGGGCACACGACCGCGCAGCCAGTTCACACACCGCCAGCGC.... 506

83 luAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99

507CGGTACTCTGAACGCGCAACGTCGCTACAAACAACGCTGGG 545

100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValSe 112

546 CACCGCGCGCATCTACGAGCTGAAGCCACACCGCTTCGTGCT..... 590

112 rThrAspProProHisHisAspHis 121

591GAACAACGACACCATCGCCACCAT 614

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-469-334-18

seq_documentation_block:

Sequence 18, Application US/08469334

Patent No. 5990383

GENERAL INFORMATION:

APPLICANT: Warren, Gregory W

APPLICANT: Koziel, Michael G

APPLICANT: Mullins, Martha A

APPLICANT: Nye, Gordon J

APPLICANT: Carr, Brian

APPLICANT: Desai, Nalini M

APPLICANT: Kostichka, N. Kristy

APPLICANT: Duck, Nicholas B

APPLICANT: Estruch, Juan J

TITLE OF INVENTION: No. 5990383el Pesticidal Proteins and Strains

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: CIBA-GEIGY Corporation

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: NY

COUNTRY: USA

ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30B

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,334

FILING DATE: 06-JUN-1995

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/463,483

FILING DATE:

APPLICATION NUMBER: US 08/314,594

FILING DATE: 09-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/218,018

FILING DATE: 23-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/037,057

FILING DATE: 25-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Spruill, W. Murray

REGISTRATION NUMBER: 32,943

REFERENCE/DOCKET NUMBER: CGC 1695/CIP3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8615

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 2004 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: misc_feature

LOCATION: 1..2004

OTHER INFORMATION: /note= "Maize optimized DNA

US-08-469-334-18

alignment_scores:

Quality: 87.00 Length: 126

Ratio: 1.338 Gaps: 6

Percent Similarity: 51.587 Percent Identity: 27.778

alignment_block:

US-09-327-750D-31 x US-08-469-334-18 ..

Align seg 1/1 to: US-08-469-334-18 from: 1 to: 11

15 GluAsnValAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAl 31

270 GAACGTGAGCATGGAGAGGTGATCTGTAGCCCAACGAGAACCTGAGCA 319

31 aAsnLysGlyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysv 48

320 ACAGCGTGGAGAGCCACTCGAGCACCACCACTGGAGGTACACCAACAC. 365

48 alProArgGlyAsnArgArgArgPheArgValArgGlnProIleLeuGln 64

366CGAGGGCGCCAGCGTGGAGCGCGCATCGGTCCCAAGGGCATCAG 410

65 TyrArgTIPAsp..... 68

411 CTTCCGCGTGAGCGTCACTACCAGCAGCAGCGAGCCGTGGCCCGCAGGAT 460

69IleMetHisArgLeuGlyGluPro...GlnAlaArgMetArgGluG 83

461 GGGGCACACGACCGCGCAGCCAGTTCACACACCGCCAGCGC.... 506

83 luAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99

507CGGTACTCTGAACGCGCAACGTCGCTACAAACAACGCTGGG 545

546 CACCGGGCCATCTAGACGTGAAGCCACCCAGCTTGTGCT..... 590

112 rThrAspProHisHisHisHisHis 121

591GAAACAGACACCATCGCGCCACCAT 614

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-463-483A-18

seq_documentation_block:

; Sequence 18, Application US/08463483A

; Patent No. 5849870

; GENERAL INFORMATION:

; APPLICANT: Warren, Gregory W

; APPLICANT: Koziel, Michael G

; APPLICANT: Mullins, Martha A

; APPLICANT: Nye, Gordon J

; APPLICANT: Carr, Brian

; APPLICANT: Desai, Nalini M

; APPLICANT: Kostichka, N. Kristy

; APPLICANT: Duck, Nicholas B

; APPLICANT: Estruch, Juan J

; TITLE OF INVENTION: No. 5849870el Pesticidal Proteins and Strains

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CIBA-GEIGY Corporation

; STREET: 7 Skyline Drive

; CITY: Hawthorne

; STATE: NY

; COUNTRY: USA

; ZIP: 10532

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30B

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/463.483A

; FILING DATE:

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/314,594

; FILING DATE: 09-SEP-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/218,018

; FILING DATE: 23-MAR-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/037,057

; FILING DATE: 25-MAR-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Spruill, W. Murray

; REGISTRATION NUMBER: 32,943

; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 919-541-8615

; TELEFAX: 919-541-8689

; INFORMATION FOR SEQ ID NO: 18:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2004 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1..2004

; OTHER INFORMATION: /note= "Maize optimized DNA

; OTHER INFORMATION: sequence for VIPIA(a) 80 kd protein from AB78"

US-08-463-483A-18

alignment_scores:

Quality: 87.00 Length: 126

Ratio: 1.338 Gaps: 6

Percent Similarity: 51.587 Percent Identity: 27.778

alignment_block:

US-09-327-750D-31 x US-08-463-483A-18 ..

Align seg 1/1 to: US-08-463-483A-18 from: 1 to: 2004

15 GluAsnValAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAl 31

||||: ||: ||||| ||: ||||| |||

270 GAACGTGAGCATGGAGAGGTGATCTGTAGCCCAACGAGAACCTGAGCA 319

||||: ||: ||||| ||: ||||| |||

31 aAsnLysGlyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysV 48

||||: ||: ||||| ||: ||||| |||

320 ACAGCGTGGAGGCCACTCGAGCCCAACTGGAGCTACACCAACAC... 365

||||: ||: ||||| ||: ||||| |||

48 alProArgGlyAsnArgArgPheArgValArgGlnProIleLeuGln 64

||||: ||: ||||| ||: ||||| |||

366CGAGGGCCAGCGTGGAGCGCGCATCGGTCCCAAGGCGATCAG 410

||||: ||: ||||| ||: ||||| |||

65 TyrArgTrpAsp..... 68

||| ||: |||

411 CTTCCGCGTGAGCGTGAACTACCGACGACGAGCCGTGGCCAGGAGT 460

||||: ||: ||||| ||: ||||| |||

69IleMetHisArgLeuGlyGluPro...GlnAlaArgMetArgGluG 83

||||: ||: ||||| ||: ||||| |||

461 GGGGCACCGACCGCCCAACACCGCCAGTTCAACACCGCCGCGC... 506

||||: ||: ||||| ||: ||||| |||

83 LuAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99

||||: ||: ||||| ||: ||||| |||

507CGGTACCTGAACGCCAACGTCGCTACACCAACCGTGGG 545

||||: ||: ||||| ||: ||||| |||

100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValSe 112

||||: ||: ||||| ||: ||||| |||

546 CACCGCGCGCATCTACGAGCTGAAGCCACCGCCAGCTTCGTGCT.... 590

||||: ||: ||||| ||: ||||| |||

112 rThrAspProHisHisHisHisHis 121

||||: ||: ||||| ||: ||||| |||

591GAAACAGACACCATCGCCACCAT 614

||||: ||: ||||| ||: ||||| |||

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-046A-18

seq_documentation_block:

; Sequence 18, Application US/08471046A

; Patent No. 5866326

; GENERAL INFORMATION:

; APPLICANT: Warren, Gregory W

; APPLICANT: Koziel, Michael G

; APPLICANT: Mullins, Martha A

; APPLICANT: Nye, Gordon J

; APPLICANT: Carr, Brian

; APPLICANT: Desai, Nalini M

; APPLICANT: Kostichka, N. Kristy

; APPLICANT: Duck, Nicholas B

; APPLICANT: Estruch, Juan J

; TITLE OF INVENTION: Method For Isolating Vegetative Insecticidal

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: NO. 5866326artis Corporation

; STREET: 3054 Cornwallis Road

; CITY: Research Triangle Park

; STATE: NC

; COUNTRY: USA

; ZIP: 27709

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30B

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/471.046A


```

100 ArgGluLysGlnLeu.....SerHisSerLeuArgAlaValSe 112
      ::::~::~:~::|||

```

OM of: US-09-327-750D-31 to: Issued_Patents_NA:* out_format : pfs
Date: Mar 11, 2002 3:50 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=framet.p2n_model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09327750/runat_11032002_101154_20340/app_query.fasta_1.1472
-DB=Issued_Patents_NA -Qfmt=fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09327750_@CGN1_1.99 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-327-750D-31
Query length: 128
Database: Issued_Patents_NA:*
Database sequences: 351203
Database length: 113238999
Search time (sec): 146.090000

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-471-033-18 +		87.00	158.56	0.3508	2004
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-044-18 +		87.00	158.56	0.3508	2004
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-463-483A-18 +		87.00	158.56	0.3508	2004
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-046A-17 +		87.00	158.56	0.3508	2004
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-046A-18 +		87.00	158.56	0.3508	2004
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-046A-17 +		87.00	158.56	0.3508	2004
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-469-334-18 +		87.00	158.56	0.3508	2004
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-469-334-18 +		87.00	158.56	0.3508	2004
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-300-529-18 +		87.00	158.56	0.3508	2004
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-471-033-35 +		87.00	155.90	0.4929	2576
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-463-483A-35 +		87.00	155.90	0.4929	2576
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-033-26 +		87.00	155.58	0.5135	2655
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-033-26 +		87.00	155.58	0.5135	2655
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-044-17 +		87.00	155.58	0.5135	2655
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-044-26 +		87.00	155.58	0.5135	2655
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-463-483A-17 +		87.00	155.58	0.5135	2655
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-463-483A-26 +		87.00	155.58	0.5135	2655
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-046A-17 +		87.00	155.58	0.5135	2655
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-046A-26 +		87.00	155.58	0.5135	2655
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-470-566B-17 +		87.00	155.58	0.5135	2655
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-470-566B-26 +		87.00	155.58	0.5135	2655
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-469-334-17 +		87.00	155.58	0.5135	2655
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-469-334-26 +		87.00	155.58	0.5135	2655
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-300-529-17 +		87.00	155.58	0.5135	2655
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-300-529-17 +		87.00	155.58	0.5135	2655
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-09-300-529-26 +		87.00	155.58	0.5135	2655
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-471-033-49 +		87.00	151.17	0.9042	4031
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-044-49 +		87.00	151.17	0.9042	4031
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-463-483A-49 +		87.00	151.17	0.9042	4031
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-471-046A-49 +		87.00	151.17	0.9042	4031
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-470-566B-49 +		87.00	151.17	0.9042	4031
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-469-334-49 +		87.00	151.17	0.9042	4031
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-147-777-3 +		86.00	129.75	14.12	24979
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-452-872-3 +		86.00	129.75	14.12	24979
/cgn2_6/ptodata/2/ina/5B_COMB.seq:PCT-US93-03985-3 +		86.00	129.75	14.12	24979
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-306-691B-23 +		79.00	139.85	3.86	2301
/cgn2_6/ptodata/2/ina/5B_COMB.seq:PCT-US93-06251-78 +		79.00	139.85	3.86	2301

/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-513-974B-31 - 78.50 146.47 1.65
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-776-971-26 - 78.50 146.47 1.65
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-513-974B-32 - 78.50 144.55 2.11
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-776-971-103 - 78.50 144.55 2.11
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-998-416-141 - 78.00 150.69 0.9614
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-998-416-141 - 78.00 150.69 0.9614

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-471-033-18

seq_documentation_block:
; Sequence 18, Application US/08471033
; Patent No. 5770696
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Kozziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,033
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: P-40,403
; REFERENCE/DOCKET NUMBER: C0C 1695/CIP3/DIV7 - SOLV3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2004 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2004
; OTHER INFORMATION: /note= "Maize optimized DNA
; OTHER INFORMATION: sequence for VIPIA(a) 80 kd protein from AB78"
US-08-471-033-18

alignment_scores:

```

|||||
212 ATGGAGTCCAAAGAGGAGGAGCGTTTAAACAATCTCATCGTGGAAAATGT 261
17 lAsnGlnGluAsnAspGluLysAspGluGlnValAlaAsnLysG 34
|||||
262 CAACCAAGGAAAATGATGAAAAAGATGAAAAGGAGCAAGTTGCTAATAAAG 311
34 lyluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
312 GGGAGCCCTTGGCCCTTACCTTTGAATGTTAGTGAATACTGTGTGCCTAGA 361
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
|||||
362 GGAACCGTAGGCGGTTCCGCGTTAGGCAGCCCATCTGCAGTATAGATG 411
67 PaspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
|||||
412 GGACATAATGCATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGGAGA 461
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
|||||
462 ATATGAAAGGATTGGGAGGAGGTGAGACAGCTGATGGAAGCTGAGG 511
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||
512 GAAAGCAGTTGAGTCATAGTTTGGCGGAGTCAGCACCTGATCCCCCTCA 561
117 sHisAspHisAspGluPheCysLeuMetPro 128
|||||
562 CCATGACCATCAGATGAGTTTGCCTTATGCCCT 595

```

seq_name: gb_est1:BE266012

seq_documentation_block:

```

LOCUS BE266012 610 bp mRNA EST 13-JUL-2000
DEFINITION 601194731F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3538735 5',
mRNA sequence.
ACCESSION BE266012
VERSION BE266012.1 GI:9139583
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Plate: L1CM224 row: m column: 08
High quality sequence stop: 607.

```

FEATURES

source

```

1..610
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3538735"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site:1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
170 a 128 c 193 g 119 t

```

BASE COUNT
ORIGIN

```

alignment_scores:
  Quality: 663.00      Length: 129
  Ratio: 5.180        Gaps: 1
  Percent Similarity: 99.225  Percent Identity: 98.450

alignment_block:
US-09-327-750D-31 x BE266012 ..
Align seg 1/1 to: BE266012 from: 1 to: 610

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
172 ATGGAGTCCAAAGAGGAGGAGCGTTTAAACAATCTCATCGTGGAAAATGT 221
17 lAsnGlnGluAsnAspGluLysAspGluGlnValAlaAsnLysG 34
|||||
222 CAACCAAGGAAAATGATGAAAAAGATGAAAAGGAGCAAGTTGCTAATAAAG 271
34 lyluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
272 GGGAGCCCTTGGCCCTTACCTTTGAATGTTAGTGAATACTGTGTGCCTAGA 321
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
|||||
322 GGAACCGTAGGCGGTTCCGCGTTAGGCAGCCCATCTGCAGTATAGATG 371
67 PaspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
|||||
372 GGACATAATGCATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGGAGA 421
84 snMetGlu.ArgIleGlyGluGluValArgGlnLeuMetGluLysLeuAr 100
|||||
422 ATATGCAACAGGATTGGGAGGAGGTGAGACAGCTGATGGAAGCTGAG 471
100 gLyuLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 117
|||||
472 GGAAGCAGTTGAGTCATAGTTTGGCGGAGTCAGCACCTGATCCCCCTC 521
117 ishIsAspHisHisAspGluPheCysLeuMetPro 128
|||||
522 ACCATGACCATCAGATGAGTTTGCCTTATGCCCC 556

```

```

VERSION      AI291126.1  GI:39333900
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 669)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL       Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Emmert-Buck, M.D., Ph.D.
             CDNA Library Preparation: M. Bento Soares, Ph.D.
             DNA Sequencing by: Greg Lennon, Ph.D.
             CDNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             www-bio.llnl.gov/bbrp/image/image.html
             Insert Length: 749 Std Error: 0.00
             Seq primer: -400P from Gibco
             High quality sequence stop: 466.
FEATURES     Location/Qualifiers
             1..669
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone="IMAGE:1881915"
             /clone_lib="NCI-CGAP_Lu5"
             /tissue_type="carcinoid"
             /lab_host="DH10B"
             /note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
             modified polylinker; 1st strand cDNA was prepared from
             neuroendocrine lung carcinoid, and was then primed with a
             Not I - oligo(dT) primer. Double-stranded cDNA was ligated
             to Eco RI adaptors (Pharmacia), digested with Not I and
             cloned into the Not I and Eco RI sites of the modified
             pT7n3 vector. Library is normalized. Library was
             constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT   186 a 132 c 185 g 164 t 2 others
ORIGIN
alignment_scores:
  Quality: 664.00      Length: 128
  Ratio: 5.270         Gaps: 0
  Percent Similarity: 98.438  Percent Identity: 98.438
alignment_block:
  US-09-327-750D-31 x AI291126
  Align seg 1/1 to: AI291126 from: 1 to: 669
1 MetGluSerLysGluGluArgAlaLeuAsnLeuIleValGluAsnVa 17
83 ATGGAGTCCAAAGACGACGCGTTAAACATCTCATCTCGTGGAAATGT 132
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
133 CAACCAAGGAAATATGATGAAAAAGATGAAAGGAGCAAGTTGCTAATAAAG 182
34 lYcGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
183 GGGAGCCCTTGGCCCTTACCTTGAATGTTAGTGAATCTGTGTGCTAGTA 232
51 GlyAsnArgArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
233 GGAAACCGTAGCGGTTCGCGCTAGGCAGCCCATCTGCATATAGATG 282
67 pAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluGluA 84
283 GGACATAATGCTAGGCTTGGAGAGCCACAGGCAAGGATGAGAGAGAGA 332

```

```

84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
|||||
333 ATATGGAAGAGATTTCGGAGGAGGTGAGACAGCTGATGGAAAAGCTGAGG 382
|||||
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHl 117
|||||
383 GAAAGCAGTTGAGTCATAGTCGCGGCGAGTCAGCACATGATCCCCCTCA 432
|||||
117 SHISAspHisHisAspGluPheCysLeuMetPro 128
|||||
433 CCATGACCATCAGCATGAGTTTGTGCTTATGCCC 466
seq_name: gb_est2:BF237433
seq_documentation_block:
LOCUS      BF237433          711 bp      mRNA          EST          14-NOV-2000
DEFINITION 601842108F1 NIH_MGC_46 Homo sapiens cDNA clone IMAGE:4079856 5',
            mRNA sequence.
ACCESSION  BF237433
VERSION    BF237433.1  GI:11151351
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 711)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: Ling Hong/Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LICM938 row: P column: 01
            High quality sequence stop: 708.
FEATURES   Location/Qualifiers
            1..711
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4079856"
            /clone_lib="NIH_MGC_46"
            /tissue_type="leiomyosarcoma cell line"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally cloned
            into EcoRI/XhoI sites using the following 5' adaptor:
            GGCACGAG(G). Size-selected >500bp for average insert size
            1.8kb. Library constructed by Ling Hong in the laboratory
            of Gerald M. Rubin (University of California, Berkeley)
            using ZAP-cDNA synthesis kit (Stratagene) and Superscript
            II RT (Life Technologies). Note: this is a NIH_MGC
            Library."
BASE COUNT 184 a 152 c 218 g 157 t
ORIGIN
alignment_scores:
  Quality: 664.00      Length: 128
  Ratio: 5.228         Gaps: 0
  Percent Similarity: 99.219  Percent Identity: 98.438
alignment_block:
  US-09-327-750D-31 x BF237433
  Align seg 1/1 to: BF237433 from: 1 to: 711
1 MetGluSerLysGluGluArgAlaLeuAsnLeuIleValGluAsnVa 17

```

```

alignment_scores:
  Quality: 668.00      Length: 128
  Ratio: 5.260         Gaps: 0
  Percent Similarity: 99.219      Percent Identity: 99.219

alignment_block:
US-09-327-750D-31 x BG705843
Align seg 1/1 to: BG705843 from: 1 to: 813.

1 MetGluSerLySGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
218 ATGGAGTCCAAAGAGGAGGAGCGTTAAACAATCTCATCGTGGAAAATGT 267
17 lAsnGlnGluAsnAspGluLysGluGlnValAlaAlaAsnLysG 34
268 CAACCAAGGAAATGATGAAAAAGATGAAAGGAGCAAGTTGCTAATAAAG 317
34 lyluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
318 GGGAGCCCTTGGCCCTACCTTGAATGTTAGTGAATCTGTGTCCTAGA 367
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
368 GGAACCCGTAGGCGGTTCCGCGTTAGGCAGCCCATCTGCAGTATAGATG 417
67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
418 GGACATATGATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGAGGAGA 467
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
468 ATATGGAAGAGTGGGAGGAGGTGACAGACGCTGATGGAAGAGCTGAGG 517
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
518 GAAAGCAGTTGAGTCATAGTCTGCGGCGAGTCAGCACTGATCCCTCA 567
117 sHisAspHisAspGluPheCysLeuMetPro 128
568 CCATGACCATCAGATGAGTTTGGCCTTATGCC 601

seq_name: gb_estl:AI291270

seq_documentation_block:
LOCUS AI291270 511 bp mRNA EST 29-JAN-1999
DEFINITION qm16e06.x1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1882018 3'
similar to SW:HG74_HUMAN Q00994 OVARIAN GRANULOSA CELL 13.0 KD
PROTEIN HGR74, ;, mRNA sequence.
ACCESSION AI291270
VERSION AI291270.1 GI:3934044
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 511)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.lnl.gov/bbrp/image/image.html

```

Insert Length: 742 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 475.
Location/Qualifiers
source
1..511
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1882018"
/clone_lib="NCI_CGAP_Lu5"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia) digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 152 a 103 c 155 g 100 t 1 others
ORIGIN

alignment_scores:
Quality: 666.00 Length: 128
Ratio: 5.286 Gaps: 0
Percent Similarity: 98.438 Percent Identity: 98.438

alignment_block:
US-09-327-750D-31 x AI291270
Align seg 1/1 to: AI291270 from: 1 to: 511

1 MetGluSerLySGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
83 ATGGAGTCCAAAGAGGAGGAGCGTTAAACAATCTCATCGTGGAAAATGT 132
17 lAsnGlnGluAsnAspGluLysGluGlnValAlaAlaAsnLysG 34
133 CAACCAAGGAAATGATGAAAAAGATGAAAGGAGCAAGTTGCTAATAAAG 182
34 lyluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
183 GGGAGCCCTTGGCCCTACCTTGAATGTTAGTGAATCTGTGTCCTAGA 232
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
233 GGAACCCGTAGGCGGTTCCGCGTTAGGCAGCCCATCTGCAGTATAGATG 282
67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
283 GGACATATGATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGAGAGA 332
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
333 ATATGGAAGAGTGGGAGGAGGTGAGAGCGCTGATGGAAGAGCTGAGG 382
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
383 GAAAGCAGTTGAGTCATAGTCTGCGGCGAGTCAGCACTGATCCCTCA 432
117 sHisAspHisAspGluPheCysLeuMetPro 128
433 CCATGACCATCAGATGAGTTTGGCCTTATGCC 466

seq_name: gb_estl:AI291126
seq_documentation_block:
LOCUS AI291126 669 bp mRNA EST 29-JAN-1999
DEFINITION qm15f02.x1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1881915 3'
similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74
(HUMAN);, mRNA sequence.
ACCESSION AI291126

277 GAAAGCAGTTGAGTCATAGTCTGCGGCGAGTCAGCACTGATCCCCCTCA 228

117 shIsAspHisAspGluPheCysLeuMet 127

227 CCATGACCATCAGATGAGTTTGCCTTATG 197

seq_name: gb_est2:BG707734

seq_documentation_block: 813 bp mRNA EST 07-MAY-2001
LOCUS BG707734 60267125F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4794061 5',
DEFINITION mRNA sequence.

ACCESSION BG707734.1 GI:13984377

VERSION BG707734

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 817)

TITLE NIH-MGC http://mgi.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cga@rs-re@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM10674 row: n column: 14

High quality sequence stop: 803.

Location/Qualifiers

1..817

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4794061"

/tissue_type="NIH_MGC_96"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescriptR (modified

pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag

); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',

size-selected for average insert size 2.3 kb and

normalized to ROT 5. This is a primary library enriched

for full-length clones and constructed using the

Cap-trapper method (Carninci, in preparation). Library

constructed by M. Brownstein (NIH/NHGRI, National

Institutes of Health). Note: this is a NIH_MGC Library."

213 a 171 c 241 g 191 t 1 others

BASE COUNT

ORIGIN

alignment_scores:

Quality: 672.00 Length: 128

Ratio: 5.250 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 99.219

alignment_block:

US-09-327-750D-31 x BG707734 ..

Align seg 1/1 to: BG707734 from: 1 to: 817

1 MetGluSerLysGluGluAlaLeuAsnValGluAsnVa 17

|||||

221 ATGGAGTCCAAAGAGACGCGTGAACATCTCATCGTGAAGAATGT 270

|||||

17 IasnGlnGluAsnAspGluLysAspGluGlnValAlaAsnLysG 34

|||||

271 CAACAGGAAATGATGAAAGATGAAAGGAGCAGTTCGCTAATAAG 320

34 LyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProAtg 50

|||||

321 GGGAGCCCTTGGCCCTTACCTTTGAATGTTAGTGAATACTGTGTGCCTAGA 370

|||||

51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67

|||||

371 GGAACCGTAGCGGTTCGCGTTAGGCAGCCCATCTGCAGTATATAGATG 420

|||||

67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84

|||||

421 GGACATAAATGCATAGCTTGGAGAGCCACAGGCAAGGATGAGAGAGAGA 470

|||||

84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100

|||||

471 ATATGGAAGAGATTGGGAGGAGGTGACAGCTGATGGAAGAGCTGAGG 520

|||||

101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 117

|||||

521 GAAAGCAGCTTGCATAGTCTGCGGCGAGTCAGCACTGATCCCCCTCA 570

|||||

117 shIsAspHisAspGluPheCysLeuMetPro 128

|||||

571 CCATGACAAATCAGATGAGTTTGCCTTATGCC 604

|||||

seq_name: gb_est2:BG705843

seq_documentation_block: 813 bp mRNA EST 07-MAY-2001

LOCUS BG705843 602669329F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4792038 5',

DEFINITION mRNA sequence.

ACCESSION BG705843

VERSION BG705843.1 GI:13980593

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 813)

TITLE NIH-MGC http://mgi.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cga@rs-re@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki

Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM10669 row: j column: 07

High quality sequence stop: 772.

Location/Qualifiers

1..813

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4792038"

/tissue_type="NIH_MGC_96"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescriptR (modified

pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag

); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',

size-selected for average insert size 2.3 kb and

normalized to ROT 5. This is a primary library enriched

for full-length clones and constructed using the

Cap-trapper method (Carninci, in preparation). Library

constructed by M. Brownstein (NIH/NHGRI, National

Institutes of Health). Note: this is a NIH_MGC Library."

207 a 168 c 243 g 194 t 1 others

BASE COUNT

ORIGIN

/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 181 a 159 c 226 g 160 t
ORIGIN

alignment_scores:
Quality: 672.00 Length: 127
Ratio: 5.291 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-327-750D-31 x BG714974 ..
Align seg 1/1 to: BG714974 from: 1 to: 726
1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuLeuValGluAsnVa 17
|||||
219 ATGGAGTCCAAAGAGGACGCGTTAAACAATCTCATCGTGGAAATGT 268
|||||
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||
269 CAACCAAGAAATCATGAAAAAGATGAAAGGAGCAAGTTGCTAATAAAG 318
|||||
34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
319 GGGAGCCCTTGGCCCTACCTTTGAATGTTAGTGAATCTGTCGCTAGA 368
|||||
51 GlyAsnArgArgPheArgValArgGlnProLeuGlnTyrArgTr 67
|||||
369 GGAACCGTAGGGGTTCCCGTTAGGCAGCCCATCTCGCAGTATAGTG 418
|||||
67 pAspIleMethHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
|||||
419 GGACATAATGCATAGGCTTGAGAGCCACAGGCAAGGATGAGAGGAGA 468
|||||
84 snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 100
|||||
469 ATATGGAAGAGATGGGGAGGAGGTGAGACAGCTGATGGAAGAGCTGAGG 518
|||||
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||
519 GAAAGACGTTGAGTCATAGTCTGCGGGCAGCTAGCACTGATCCCTCA 568
|||||
117 sHisAspHisAspGluPheCysLeuMet 127
|||||
569 CCATGACCATCAGCATGAGTTTGGCCTATG 599
|||||

seq_name: gb_est2:BF967675
seq_documentation_block:
LOCUS BF967675 793 bp mRNA EST 23-JAN-2001
DEFINITION 602287383T1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4374549 3',
mRNA sequence.
ACCESSION BF967675
VERSION BF967675.2 GI:12388141
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 793)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT On Jan 16, 2001 this sequence version replaced gi:12334890.
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10038 row: f column: 22
High quality sequence start: 20
High quality sequence stop: 762.
FEATURES
Location/Qualifiers
1..793
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4374549"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 185 a 235 c 170 g 203 t
ORIGIN
alignment_scores:
Quality: 672.00 Length: 127
Ratio: 5.291 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-327-750D-31 x BF967675/rev ..
Align seg 1/1 to reverse of: BF967675 from: 1 to: 793
1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuLeuValGluAsnVa 17
|||||
577 ATGGAGTCCAAAGAGGAGGCGTTAAACAATCTCATCGTGGAAATGT 528
|||||
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||
527 CAACCAAGAAATCATGAAAAAGATGAAAGGAGCAAGTTGCTAATAAAG 478
|||||
34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
477 GGGAGCCCTTGGCCCTACCTTTGAATGTTAGTGAATCTGTCGCTAGA 428
|||||
51 GlyAsnArgArgPheArgValArgGlnProLeuGlnTyrArgTr 67
|||||
427 GGAACCGTAGGGGTTCCCGTTAGGCAGCCCATCTCGCAGTATAGTG 378
|||||
67 pAspIleMethHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
|||||
377 GGACATAATGCATAGGCTTGAGAGCCACAGGCAAGGATGAGAGGAGA 328
|||||
84 snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 100
|||||
327 ATATGGAAGAGATGGGGAGGAGGTGAGACAGCTGATGGAAGAGCTGAGG 278
|||||
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||

Percent Similarity: 100.000 Percent Identity: 99.219

alignment_block:
US-09-327-750D-31 x BG285825

Align seg 1/1 to: BG285825 from: 1 to: 946

```

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
190 ATGGAGTCCAAAGAGGAGGAGCGTTAAACAATCTCATCTGCGGAAAATGT 239
17 IasnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||
240 CAACACGAGAAATGATGAAAAAGATGAAAGAGGACCAAGTTGCTAATAAG 289
34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
290 GGGAGCCCTTTGGCCCTTACCTTTGATGTTAGTGAATTAATCTGTCCTAGA 339
51 GlysAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
|||||
340 GGAACCGTAGCGGTCCCGTTAGGACGCCCATCTCTGCAGTATAGATG 389
67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
|||||
390 GGACATAATGCATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGAGGAGA 439
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
|||||
440 ATATGGAAGAGGATCGGGAGGAGGTGAGACAGCTGTATGGAAGAGCTGAG 489
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||
490 GAAACGAGTTGAGTCACTAGTCTGCGGGCAGTCAGCACTGATCCCCCTCA 539
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
540 CCATGACCATCAGCATGAGTTTGCCTTATGCCCC 573

```

seq_name: gb_est2:BG714974

seq_documentation_block:
LOCUS BG714974 726 bp mRNA 08-MAY-2001
DEFINITION 602677263F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4799823 5',
mRNA sequence.
ACCESSION BG714974
VERSION BG714974.1 GI:13993905
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 726)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10689 row: n column: 16
High quality sequence stop: 723.
Location/Qualifiers
1..726
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4799823"

FEATURES
source

seq_documentation_block:
LOCUS BG285825 946 bp mRNA 21-FEB-2001
DEFINITION 602380917F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4498445 5',
mRNA sequence.
ACCESSION BG285825
VERSION BG285825.1 GI:13038171
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 946)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10361 row: a column: 06
High quality sequence stop: 710.
Location/Qualifiers
1..946
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4498445"
/clone_lib="NIH_MGC_93"
/lab_host="DH10B (phage-resistant)"
/lab_type="transitional cell papilloma, cell line"
/note="Organ: bladder; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
276 a 201 c 279 g 190 t
BASE COUNT
ORIGIN

alignment_scores:
Quality: 674.00 Length: 128
Ratio: 5.266 Gaps: 0


```
VERSION BG707398.1 GI:13983707
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10678 row: d column: 04
High quality sequence stop: 814.

FEATURES
source
Location/Qualifiers
1..835
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4795347"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
size-selected for average insert size 2.3 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 220 a 176 c 246 g 193 t
ORIGIN

alignment_scores
Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block
US-09-327-750D-31 x BG707398 ...
Align seg 1/1 to: BG707398 from: 1 to: 835

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
248 ATGGAGTCCAAAGAGAGAGAGAGGCGTTAAACAATCTCATCGTGAAAAATGT 297
17 lAsnGlnGluAsnAspGluLysAspGluGlnValAlaAlaAsnLysC 34
17 lAsnGlnGluAsnAspGluLysAspGluGlnValAlaAlaAsnLysC 34
298 CAACACGAGAAATGATGAAAAAGATGAAAGGAGCAAGTTGCTAATAAAG 347
34 lylGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
348 GGGAGCCCTGGCCCTACCTTGAATGTTAGTCAATCTGTCGCTAGC 397
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
398 GGAACCGTAGGCGGTTCGCGTTAGCGACGCCCATCTCGACGTATAGATG 447
67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu 84
448 GGACATAATGATAGCTTGGAGAGGCACAGGCAAGGATGAGAGAGAGA 497
```

```
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
498 ATATGCAAGAGATTGGGAGGAGTGCAGCAGCTGATGGAAGCTGAGG 547
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
548 GAAACACAGTTGAGTCATAGTCTGCGGCGCAGTCAGTATCCCTCA 597
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
598 CCATGACCATCAGCATGAGTTTTCCTTATGCCC 631
seq_name: gb_est2:BF038787
seq_documentation_block:
LOCUS BF038787 1042 bp mRNA EST 20-OCT-2000
DEFINITION 601462119f1 NIH_MGC_66 Homo sapiens cDNA clone IMAGE:3865348 5',
mRNA sequence.
ACCESSION BF038787
VERSION BF038787.1 GI:10745987
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1042)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DP
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9608 row: f column: 05
High quality sequence stop: 732.
FEATURES
source
Location/Qualifiers
1..1042
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3865348"
/clone_lib="NIH_MGC_66"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Ovary; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies."
BASE COUNT 308 a 200 c 296 g 237 t 1 others
ORIGIN

alignment_scores
Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block
US-09-327-750D-31 x BF038787
Align seg 1/1 to: BF038787 from: 1 to: 1042

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
85 ATGGAGTCCAAAGAGAGAGAGGCGTTAAACAATCTCATCGTGAAAAATGT 134
17 lAsnGlnGluAsnAspGluLysAspGluGlnValAlaAlaAsnLysC 34
135 CAACACGAGAAATGATGAAAAAGATGAAAGGAGCAAGTTGCTAATAAAG 184
```

BASE COUNT 226 a 139 c 222 g 185 t
ORIGIN

alignment_scores:
Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-327-750D-31 x BE889313 ..
Align seg 1/1 to: BE889313 from: 1 to: 772

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
125 ATGGAGTCCAAAGAGAGACGCGTTAAACAATCTCATCTGGAATAATGT 174
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
175 CAACACGAGAAATGATGAAAAAGATGAAAGGAGCAAGTTGCTAATAAAG 224
34 lylGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
225 GGGAGCCCTTGGCCCTACCTTTGAATGTTAGTGAATCTGTGTGCCTAGA 274
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
275 GGAACCCGTAGCGGTTCCGCGTTAGGCAGCCCATCTCGCATATAGATG 324
67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu 84
325 GGACATAATGCATAGCTTGGAGACCCACAGGCAAGGATCAGAGAGGAGA 374
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
375 ATATGGAAGAGGATGGGAGGAGGTGAGACAGCTGATGGAAGAGCTGAGG 424
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
425 GAAAGACAGTTGAGTCATAGTTTGGCGGCAGTCAGACACTGTATCCCTCA 474
117 shiAspHisHisAspGluPheCysLeuMetPro 128
475 CCATGACCATCAGATGAGTTTGGCTTATGCCCC 508

seq_name: gb_est2:BG709503

seq_documentation_block:
LOCUS BG709503 810 bp mRNA EST 07-MAY-2001
DEFINITION 602674781F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4797213 5',
mRNA sequence.
ACCESSION BG709503
VERSION BG709503.1 GI:13987904
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 810)
AUTHORS cDNA Library Preparation: Miklos Palkovits, M.D., Ph.D.
Toshiyuki and Piero Carninci (RIKEN)
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov
Plate: LLAM10683 row: a column: 22
High quality sequence stop: 770.
FEATURES
Location/Qualifiers
source
1..810

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4797213"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 206 a 168 c 247 g 187 t
ORIGIN

alignment_scores:
Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-31 x BG709503 ..

Align seg 1/1 to: BG709503 from: 1 to: 810

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
218 ATGGAGTCCAAAGAGAGACGCGTTAAACAATCTCATCTGGAATAATGT 267
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
268 CAACACGAGAAATGATGAAAAAGATGAGAGGCAAGTTGCTAATAAAG 317
34 lylGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
318 GGGAGCCCTTGGCCCTACCTTTGAATGTTAGTGAATCTGTGTGCCTAGA 367
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
368 GGAACCCGTAGCGGTTCCGCGTTAGGCAGCCCATCTCGCATATAGATG 417
67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu 84
418 GGACATAATGCATAGCTTGGAGACCCACAGGCAAGGATCAGAGAGGAGA 467
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
468 ATATGGAAGAGGATGGGAGGAGGTGAGACAGCTGATGGAAGAGCTGAGG 517
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
518 GAAAGACAGTTGAGTCATAGTTTGGCGGCAGTCAGACACTGTATCCCTCA 567
117 shiAspHisHisAspGluPheCysLeuMetPro 128
568 CCATGACCATCAGATGAGTTTGGCTTATGCCCC 601

seq_name: gb_est2:BG707398

seq_documentation_block:
LOCUS BG707398 835 bp mRNA EST 07-MAY-2001
DEFINITION 602672810F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4795347 5',
mRNA sequence.
ACCESSION BG707398

241 GGGAGCCCTGGCCCTACCTTTGAATGTTAGTGAATACTGTGTGCCTAGA 290

51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67

291 GGAACCGTAGCGGTTCCCGGTAGGACGCCATCCTCGATATAGATG 340

67 PaspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu 84

341 GGACATAATGCATAGCTTGGAGAGCCACAGGCAAGGATGAGAGGAGA 390

84 snMetGluArgIleGlyGluGlnValArgGlnLeuMetGluLysLeuArg 100

391 ATATGAAAGGATGGGAGGAGGTGAGACAGCTGATGAAAGCTGAGG 440

101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117

441 GAAAGCAGTTGAGTCATAGTTGGGGCAGTCAGCAGTATGCCCTCA 490

117 shIsAspHisAspGluPheCysLeuMetPro 128

491 CCATGACCATCAGCATGAGTTTTCCTTATGCC 524

seq_name: gb_est1.AL520931

seq_documentation_block:

LOCUS AL520931 717 bp mRNA EST 13-FEB-2001
DEFINITION AL520931 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB002YA15 3
prime. mRNA sequence.

ACCESSION AL520931

VERSION AL520931.1 GI:12784424

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 717)

AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1..717
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB002YA15"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 157 a 217 c 153 g 188 t 2 others
ORIGIN

alignment_scores:

Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-31 x AL520931/rev ..

Align seg 1/1 to reverse of: AL520931 from: 1 to: 717

1 MetGluSerLysGluGlnArgAlaLeuAsnAsnLeuIleValGluAsnVa 17

530 ATGGAGTCCAAAGAGGACGAGCGTTAAACAATCTCATCGTGGAAAATGT 481

17 LAsnGlnGluAsnAspGluLysAspGluGlnValAlaAsnLysG 34

480 CAACACAGGAAATGATGAAAAAGATGAAAGGAGCAAGTTCCTAATAAAG 431

34 LylGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50

430 GGGAGCCCTTGGCCCTACCTTGAATGTTAGTGAATACTGTGTGCCTAGA 381

51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67

380 GGAACCGTAGCGGTTCCCGTTAGGACGCCATCCTGCAGTATAGATG 331

67 PaspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu 84

330 GGACATAATGCATAGCTTGGAGAGCCACAGGCAAGGATGAGAGGAGA 281

84 snMetGluArgIleGlyGluGlnValArgGlnLeuMetGluLysLeuArg 100

280 ATATGAAAGGATGGGAGGAGGTGAGACAGCTGATGGAAGAGCTGAGG 231

101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117

230 GAAAGCAGTTGAGTCATAGTCTCGCGGCGAGTGCAGCTGATCCCCCTCA 181

117 shIsAspHisAspGluPheCysLeuMetPro 128

180 CCATGACCATCAGCATGAGTTTTCCTTATGCC 147

seq_name: gb_est2:BE889313

seq_documentation_block:

LOCUS BE889313 772 bp mRNA EST 20-OCT-2000
DEFINITION 601513201F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3914401 5',
mRNA sequence.

ACCESSION BE889313

VERSION BE889313.1 GI:10346503

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 772)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-re@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Arrayed by: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LHAM9736 row: b column: 02

High quality sequence stop: 763.

Location/Qualifiers

1..772

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3914401"

/clone_lib="NIH_MGC_71"

/tissue_type="leiomyosarcoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site:1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

OM of: US-09-327-750D-31 to: EST:* out_format : pfs
Date: Mar 11, 2002 2:16 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet.p2n.model -DEV=xlh
-Q/Con2_1/USPTO.spool/US09327750/runat_11032002_101153_20308/app_query.fasta_1.1472
-DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blotsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09327750 -CGNI_1.5654
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-31
Query length: 128
Database: EST:*
Database sequences: 11351937
Database length: 1077921985
Search time (sec): 4085.940000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
9b_est1:BE314909	+	679.00	1289.59	1.1e-62	594	! BE314909 601140074F1 NIH_MGC_9
9b_est1:AL520931	+	679.00	1287.82	1.4e-62	717	! AL520931 AL520931 LTI_NFL004_NE
9b_est1:BE889313	+	679.00	1287.12	1.6e-62	772	! BE889313 601513201F1 NIH_MGC_71
9b_est2:BG709503	+	679.00	1286.67	1.7e-62	810	! BG709503 602674781F1 NIH_MGC_96
9b_est2:BG707398	+	679.00	1286.38	1.7e-62	835	! BG707398 602672810F1 NIH_MGC_96
9b_est2:BF038787	+	679.00	1284.29	2.3e-62	1042	! BF038787 601462119F1 NIH_MGC_96
9b_est2:BG285825	+	674.00	1275.65	6.9e-62	946	! BG285825 602380917F1 NIH_MGC_96
9b_est2:BG714974	+	672.00	1274.33	8.1e-62	726	! BG714974 602677263F1 NIH_MGC_96
9b_est2:BG707675	+	672.00	1273.49	9.0e-62	793	! BG707675 602287383T1 NIH_MGC_96
9b_est2:BG707734	+	672.00	1273.21	9.4e-62	817	! BG707734 602671255F1 NIH_MGC_96
9b_est2:BG705843	+	668.00	1265.62	2.5e-61	813	! BG705843 602669329F1 NIH_MGC_96
9b_est1:AI9291270	+	666.00	1266.17	2.3e-61	511	! AI9291270 qm16e06.x1 NCI_CGAP_Lu
9b_est1:AI9291126	+	664.00	1259.81	5.2e-61	669	! AI9291126 qm15f02.x1 NCI_CGAP_Lu
9b_est2:BF237433	+	664.00	1259.24	5.6e-61	711	! BF237433 601842108F1 NIH_MGC_48
9b_est1:BE733217	+	663.00	1258.77	6.0e-61	610	! BE733217 601194731F1 NIH_MGC_7
9b_est1:BE733217	+	661.00	1251.32	1.6e-60	897	! BE733217 601568452F1 NIH_MGC_21
9b_est2:BF967457	+	658.00	1247.04	2.7e-60	769	! BF967457 602287383F1 NIH_MGC_96
9b_est1:BG715745	+	658.00	1246.48	2.9e-60	816	! BG715745 602676845F1 NIH_MGC_96
9b_est2:BF685596	+	657.00	1244.55	3.7e-60	818	! BF685596 602142451F1 NIH_MGC_48
9b_est2:BG709294	+	646.00	1223.96	5.2e-59	782	! BG709294 602674689F1 NIH_MGC_96
9b_est1:AL520932	+	644.00	1219.53	9.2e-59	834	! AL520932 AL520932 LTI_NFL004_NE
9b_est2:BG715353	+	638.00	1209.80	3.2e-58	694	! BG715353 602677380F1 NIH_MGC_96
9b_est2:BG334975	+	636.00	1204.56	6.2e-58	807	! BG334975 602403408F1 NIH_MGC_21
9b_est2:BG655001	+	606.50	1151.87	5.4e-55	547	! BG655001 ib44f06.y1 HR85 islet
9b_est2:BG72167	+	606.50	1150.48	6.4e-55	634	! BG72167 602513894F1 NIH_MGC_16
9b_est2:BF971303	+	606.50	1150.13	6.7e-55	658	! BF971303 602273150F1 NIH_MGC_84
9b_est2:BG714823	+	606.50	1149.91	6.9e-55	674	! BG714823 602677146F1 NIH_MGC_96
9b_est1:AV702285	+	606.50	1149.68	7.1e-55	690	! AV702285 AV702285 ADB Homo sapi
9b_est1:AV705808	+	606.50	1149.67	7.1e-55	691	! AV705808 AV705808 ADB Homo sapi
9b_est1:AL526209	+	606.50	1149.62	7.2e-55	695	! AL526209 AL526209 LTI_NFL003_NH
9b_est2:BG820179	+	606.50	1149.56	7.2e-55	699	! BG820179 602282222F1 NCI_CGAP_Lu
9b_est1:BE790774	+	606.50	1148.49	8.3e-55	783	! BE790774 601581936F1 NIH_MGC_7
9b_est1:AL526247	+	606.50	1148.36	8.4e-55	794	! AL526247 AL526247 LTI_NFL003_NE
9b_est1:AV702643	+	603.50	1144.05	1.5e-54	686	! AV702643 AV702643 ADB Homo sapi
9b_est1:AI929703	+	599.50	1138.01	3.2e-54	576	! AI929703 au63f04.y1 Schneider f
9b_est2:BG715659	+	599.50	1135.73	4.3e-54	734	! BG715659 602676924F1 NIH_MGC_96
9b_est1:AV732454	+	596.00	1135.10	4.6e-54	386	! AV732454 AV732454 HFF Homo sapi
9b_est2:BF316327	+	596.00	1126.82	1.3e-53	929	! BF316327 601901880F1 NIH_MGC_19
9b_est1:AI141583	+	592.00	1123.54	2.0e-53	585	! AI141583 qa90g01.x1 Soares_feta
9b_est2:BG709325	+	590.50	1117.40	4.5e-53	828	! BG709325 602673345F1 NIH_MGC_96

9b_est1:AI826617 - 590.00 1119.74 3.3e-53 584 ! AI826617 wk35d03.x1 NCI_CGAP
9b_est1:AW160832 + 584.50 1110.13 1.1e-52 531 ! AW160832 au76b12.y1 Schneide
9b_est2:BF346478 + 583.50 1106.22 1.9e-52 656 ! BF346478 602020342F1 NCI_CGA
9b_est1:AV727777 + 580.50 1101.83 3.3e-52 569 ! AV727777 AV727777 HTC Homo s
9b_est1:AW953922 + 580.50 1101.23 3.6e-52 607 ! AW953922 EST365887 MAGE rese
seq_name: gb_est1:BE314909
seq_documentation_block:
LOCUS BE314909 594 bp mRNA EST 26-OCT-2000
DEFINITION 601140074F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3049290 5',
mRNA sequence.
ACCESSION BE314909
VERSION BE314909.1 GI:9145006
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/BTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM83 row: c column: 19
High quality sequence stop: 594.
Location/Qualifiers
1. 594
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3049290"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
ECORI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGCAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 161 a 126 c 181 g 126 t
ORIGIN

alignment_scores:

Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-31 x BE314909 ..

Align seg 1/1 to: BE314909 from: 1 to: 594

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
141 ATGGAGTCCAAAGAGCAAGCGCGTAAACAATCTCATCTGCGAAAATGT 190
17 IAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||
191 CAACCGAGAAATGATGAAAAAGATGAAAGGCAAGGTGCTATAAAG 240
34 IyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||

LOCUS BC003254 2901 bp mRNA 20-FEB-2001
DEFINITION Mus musculus, similar to dentatorubral pallidoluysian atrophy,
clone MGC:5758, mRNA, complete cds.
ACCESSION BC003254
VERSION BC003254.1 GI:13096906
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2901)
Direct Submission
Submitted (20-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalona@bcm.tmc.edu
Villalón, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 9 Row: 0 Column: 6.
Location/Qualifiers
1. .2901
/organism="Mus musculus"
/bb_xref="taxon:10090"
/clone="MGC:5758 IMAGE:3500522"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
755. .2359
/codon_start=1
/product="Similar to dentatorubral pallidoluysian atrophy"
/protein_id="AAH03254.1"
/bb_xref="GI:13096907"
/translation="MYPQLLPTLPGLLPLKPIIHVTHLPRAPKEHTPTSHQSLOSP
PQPLPFPFPPPAATGTPPLTATQIQEPAEEYEPPEPSPAPKSPPPKVVDPVS
HAQSARFNKHLDRGFNSCARSDLYFVLEGSKLAKRADLVKVRREAQRAREEKE
RERERERERERERERELVMSKLAQRAVECPSLGVPVHRPFPFGSAVATVP
VYLGDTPALRTLSVARHVMSPGNRHPPVYVPLGAVDPLGYNVPAIYSSDPAARE
REREARERDLRLKPGFEVKESELEPLHGVPGLDFPRHGGALQPGPGLHPFP
FHPSPGLERLRLAALGAPALRPMDSYAERLAARQHAERVAALGNDPLARLQMLNLT
PHHGHSHTHSLHLHQQDAHAASVSHYPLTIDPLASGSHLTRIPYAGTLFNPLLP
PCHNEVLRHQLFASLPYRDLPASLAPMSAAHQMQAQLQMAHQSAELQRLALEQQQWLHAH
HPLHSVPLPAQEDYYSHLKESDKPL"
BASE COUNT 623 a 958 c 791 g 529 t
ORIGIN

alignment_scores:
Quality: 384.50 Length: 119
Ratio: 3.607 Gaps: 3
Percent Similarity: 87.395 Percent Identity: 65.546

alignment_block:
US-09-327-750D-32 x BC003254

```

alignment_scores:
  Quality: 446.50      Length: 129
  Ratio: 3.917        Gaps: 3
  Percent Similarity: 88.372      Percent Identity: 68.217

alignment_block:
  US-09-327-750D-32 x AF097437
  Align seg 1/1 to: AF097437 from: 1 to: 2269

  1 MetGluSerlysluylusArgAlaValAsnSerLeuSerMetGluAsnAl 17
  1268 ATGGAGTCCAAAGAT...CAAGCGGTGAAATAATCTCAACATGGAGATGA 1314
  17 aAsnGlnGluAsnGluGluLysGlu.....GlnValAlaAsnLysG 31
  1315 CCATCAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1364
  31 lyGluPro...LeuAlaLeuProLeuAspAlaGlyGluTyrCysValPro 46
  1365 GGGAGCCAGCTGTGGCCCTGACCTCCAGGCTGGCAAAATCTGTCACCT 1414
  47 ArgGlyAsnArgArgArgPheProValArgGlnProIleLeuGlnTyrAr 63
  1415 AGAGGAGGTGCGAGCGGTTCGGGGTTCGGCAGCCATCGCTCACTATAG 1464
  63 gTrpAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluG 80
  1465 ATGGGACCTGATCCAGAGGTTCGGGAGGCCCCAGGAGGATGAGAGG 1514
  80 luAsnMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeu 96
  1515 AGAAGCTACAGAGGTTTGGGGGTGATGTGAGACAGCTCATGGAGAGCTG 1564
  97 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 113
  1565 AGGGAAGGAGGAGTGCAGCACACCTCGCGGGGTTAGCACTGACCGGCC 1614
  113 oHisHisAspHisAspGluPheCysLeuMetPro 125
  1615 TCATCATGACCACCATGATGAGTTTGCCCTCATGCCCC 1651

seq_name: gb_ro:AF097438

seq_documentation_block:
  LOCUS AF097438 835 bp mRNA ROD 11-APR-1999
  DEFINITION Mus musculus brain expressed x-linked protein 1 (Bex1) mRNA,
  complete cds.
  ACCESSION AF097438
  VERSION AF097438.1 GI:4580589
  KEYWORDS
  SOURCE
  ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  Brown,A.L. and Kay,G.F.
  Bex1, a gene with increased expression in parthenogenetic embryos,
  is a member of a novel gene family on the mouse X chromosome
  Hum. Mol. Genet. 8 (4), 611-619 (1999)
  99172070
  Erratum: [[published erratum appears in Hum Mol Genet 1999
  May; 8(5): 943]]
  2 (bases 1 to 835)
  Brown,A.L. and Kay,G.F.
  Direct Submission
  Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
  Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
  Location/Qualifiers
  1..835
  /organism="Mus musculus"
  /db_xref="taxon:10090"
  /chromosome="X"

gene
  CDS
  /map="near Plp"
  /tissue_type="blastocyst"
  1..835
  /gene="Bex1"
  205..591
  /gene="Bex1"
  /codon_start=1
  /product="brain expressed X-linked protein 1"
  /protein_id="AAD24429.1"
  /db_xref="GI:4580590"
  /translation="MESKDGQVKNNMNDHQKKEEKPODTIRRPVALISEAG
  KNCAPGRRRRFVRQPIAHYRWDLQMRVGPQGRREENVQRFQGDVQRLMEKLRR
  QLSHSLRAVSTDPHPHHDHDFCLMP"

BASE COUNT 238 a 184 c 246 g 167 t
ORIGIN

alignment_scores:
  Quality: 446.50      Length: 129
  Ratio: 3.917        Gaps: 3
  Percent Similarity: 88.372      Percent Identity: 68.217

alignment_block:
  US-09-327-750D-32 x AF097438
  Align seg 1/1 to: AF097438 from: 1 to: 835

  1 MetGluSerlysluylusArgAlaValAsnSerLeuSerMetGluAsnAl 17
  205 ATGGAGTCCAAAGAT...CAAGCGGTGAAATAATCTCAACATGGAGATGA 251
  17 aAsnGlnGluAsnGluGluLysGlu.....GlnValAlaAsnLysG 31
  252 CCATCAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 301
  31 lyGluPro...LeuAlaLeuProLeuAspAlaGlyGluTyrCysValPro 46
  302 GGGAGCCAGCTGTGGCCCTGATCTCCAGGCTGGCAAAATCTGTCGCT 351
  47 ArgGlyAsnArgArgArgPheProValArgGlnProIleLeuGlnTyrAr 63
  352 AGAGGAGGTGCGAGCGGTTCGGGGTTCGGCAGCCATCGCTCACTATAG 401
  63 gTrpAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluG 80
  402 ATGGGACCTGATCCAGAGGTTCGGGAGGCCCCAGGAGGATGAGAGG 451
  80 luAsnMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeu 96
  452 AGAAGCTACAGAGGTTTGGGGGTGATGTGAGACAGCTCATGGAGAGCTG 501
  97 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 113
  502 AGGGAAGGAGGAGTGCAGCACACCTCGCGGGGTTAGCACTGACCGGCC 551
  113 oHisHisAspHisAspGluPheCysLeuMetPro 125
  552 TCATCATGACCACCATGATGAGTTTGCCCTCATGCCCC 588

seq_name: gb_ro:AF051347

seq_documentation_block:
  LOCUS AF051347 838 bp mRNA ROD 01-OCT-1998
  DEFINITION Mus musculus REX-3 mRNA, complete cds.
  ACCESSION AF051347
  VERSION AF051347.1 GI:3510642
  KEYWORDS
  SOURCE
  ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  Faria,T.N., LaRosa,G.J., Wilen,E., Liao,J. and Gudas,L.J.

```

```

|||||
18762 CCATGACCATCAGATGAGTTTGCCTTATGCC 18795

seq_name: gb_ro:AF097439

seq_documentation_block:
LOCUS AF097439 785 bp mRNA ROD 13-APR-1999
DEFINITION Mus musculus brain expressed X-linked protein 2 (Bex2) mRNA,
complete cds.
ACCESSION AF097439
VERSION AF097439.1 GI:4580591
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Brown,A.L. and Kay,G.F.
TITLE Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
JOURNAL Hum. Mol. Genet. 8 (4), 611-619 (1999)
MEDLINE 99172070
REMARK Erratum: [[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
REFERENCE
AUTHORS Brown,A.L. and Kay,G.F.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
FEATURES
source
1..785
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="dbEST:W48832"
/chromosome="X"
/map="near Pip"
/dev_stage="embryo; 15.5 dpc"
1..785
/gene="Bex2"
139..528
/gene="Bex2"
/product="brain expressed X-linked protein 2"
/protein_id="AAD24430.1"
/db_xref="GI:4580592"
/translation="MESKVGQVKNLNMNDHQKEEKPKQDASKRDPVALPPEA
GDYVPRGRRRFRVPIVHYRWDLMRVGEPQGRMRREENVQRFQDVRQLMEKLR
QLSHSLRAVSTDPHPHHDHDFCLMP"
BASE COUNT 212 a 163 c 235 g 175 t
ORIGIN

alignment_scores:
Quality: 496.00 Length: 129
Ratio: 4.168 Gaps: 2
Percent Similarity: 92.248 Percent Identity: 72.093

alignment_block:
US-09-327-750D-32 x AF097439 ..
Align seg 1/1 to: AF097439 from: 1 to: 785

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
139 ATGGAGTCCAAAGTGGAACAAGCGGTGAAAAATCTCAACATGGAGTA 188
|||||
17 aAsnGlnGluAsnGluLysGlu.....GlnValAlaAsnLysG 31
|||||
189 CCATCAGGAAGAGGAGAAAGGACCAAGCCACAGGATGTCAGCAAA 238
|||||
31 LysGluProLeu...AlaLeuProLeuAspAlaGlyGluTyrCysValPro 46
|||||

```

```

239 GGGATCCGATTGTGGCCCTGCCTTTCCAAAGCTGGAGACTACTACGTGCCT 288
|||||
47 ArgGlyAsnArgArgPheProValArgGlnProIleLeuGlnTyrAr 63
|||||
289 AGAGGAGGTGGAGCGGTTCGGGGTTCGCAGCCCATCGTCACACTACAG 338
|||||
63 gTrpAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGlu 80
|||||
339 ATGGGACCTGCATGCATAGGTTGGGGAGCCCCAGGAGGATGAGAGAGG 388
|||||
80 LuAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 96
|||||
389 AGAACGTACAGAGGTTTGGGATGATGTGAGACAGCTCATGGAGAAGCTG 438
|||||
97 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 113
|||||
439 AGGGAAGCGACGTGAGCACAGCCCTGCGGGGGTGTAGCAGCTGACCCGCC 488
|||||
113 OhIshHisAspHisHisAspGluPheCysLeuMetPro 125
|||||
489 TCATCATGACCAACCATGATGAGTTTGCCTTATGCC 525
|||||

seq_name: gb_ro:AF097437

seq_documentation_block:
LOCUS AF097437 2269 bp DNA ROD 05-AUG-1999
DEFINITION Mus musculus Bex1 protein (Bex1) gene, complete cds.
ACCESSION AF097437
VERSION AF097437.1 GI:5702155
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Brown,A.L. and Kay,G.F.
TITLE Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
JOURNAL Hum. Mol. Genet. 8 (4), 611-619 (1999)
MEDLINE 99172070
PUBMED 10072429
REFERENCE
AUTHORS Brown,A.L. and Kay,G.F.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4030, Australia
FEATURES
source
1..2269
/organism="Mus musculus"
/strain="129SVJ"
/db_xref="taxon:10090"
/map="near Pip"
/tissue_type="liver"
/join(409..509,866..942,1263..1873)
/gene="Bex1"
/product="Bex1 protein"
409..1873
/gene="Bex1"
/note="expressed in brain; X-linked"
1268..1654
/gene="Bex1"
/codon_start=1
/product="Bex1 protein"
/protein_id="AAD47168.1"
/db_xref="GI:5702156"
/translation="MESKDGQVKNLNMNDHQKEEKPKQDTRREPAVLTSEAG
KNCAPRGRRRFRVPIAHYRWDLMRVGEPQGRMRREENVQRFQDVRQLMEKLR
QLSHSLRAVSTDPHPHHDHDFCLMP"
BASE COUNT 563 a 554 c 705 g 447 t
ORIGIN

```


/note="L1ME2 repeat: matches 4631. .6155 of consensus"
 34528. .34786
 /note="L1ME3A repeat: matches 4404. .4664 of consensus"
 34787. .35085
 /note="Alusg repeat: matches 1. .300 of consensus"
 35086. .35412

alignment_scores:
 Quality: 602.50 Length: 128
 Ratio: 4.898 Gaps: 1
 Percent Similarity: 96.094 Percent Identity: 89.062

alignment_block:

US-09-327-750D-32 x AL133348/rev ..

Align seg 1/1 to reverse of: AL133348 from: 1 to: 40584

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 |||||
 31132 ATGGAGTCCAAAGAGGAAACGAGCGTTAAACAATCTCATCTGGAATGT 31083

17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
 :|||
 31082 CAACCCAGAAATGATGAAAAAGATGAAAGGAGCAAGTTGCTAATAAG 31033

31 lyGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
 |||||
 31032 GGGAGCCCTGGCCCTACCTTTGAATTTAGTGAATACTGTGTGCCTAGA 30983

48 GlyAsnArgArgPheProValArgGlnProIleLeuGlnTyrArgTr 64
 |||||
 30982 GGAACCGTAGGCGGTTCCCGCTTAGCGAGCCATCTCGCAGTATAGT 30933

64 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 81
 |||||
 30932 GGACATAATCATAGCTTGAGAGCCACAGGCAAGATGAGAGAGAGA 30883

81 snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 97
 |||||
 30882 ATATGGAAGGATTTGGGAGGAGTGAGACAGCTGATGGAAGCTCAGG 30833

98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 114
 |||||
 30832 GAAAGCAGTTGAGTCATAGTTTCGCGGCAGTCAGCACTGATCCCCCTCA 30783

114 sHisAspHisAspGluPheCysLeuMetPro 125

30782 CCATGACCATCAGTGTGTTTGGCTTATGCCC 30749

seq_name: gb_htg:HSU80B1

seq_documentation_block:
 LOCUS HSU80B1 41029 bp DNA HTG 10-JUL-2001
 DEFINITION Homo sapiens chromosome X clone LLOXNC01-80B1 map q22.1-22.3, ***
 SEQUENCING IN PROGRESS ***, in ordered pieces.

ACCESSION AL022169
 VERSION AL022169.3 GI:13276704
 KEYWORDS HTG; HTGS_PHASE2; HTGS_CANCELLED.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 41029)

Bird.C.
 Direct Submission
 Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk

On Mar 12, 2001 this sequence version replaced gi:12750927.

----- Genome Center

Center: Sanger Centre
 Center code: SC
 Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk
 ----- Project Information
 Center project name: cu80B1
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: plasmid; L08752; 100% of reads
 Chemistry: Dye-terminator ABI; 19% of reads
 Chemistry: Dye-terminator Big Dye; 80% of reads
 Consensus quality: 41009 bases at least Q40
 Consensus quality: 41023 bases at least Q30
 Consensus quality: 41027 bases at least Q20
 Insert size: 41029; sum-of-contigs
 Quality coverage: 51751; 0.6% error; agarose-fp
 coverage: 7.67x in Q20 bases; sum-of-contigs Quality
 coverage: 7.67x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES

source

1..41029
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="X"
 /map="q22.1-22.3"
 /clone="LLOXNC01-80B1"
 /clone_lib="LLOXNC01"

misc_feature

1..41029
 /note="assembly_fragment:00890"
 BASE COUNT 13249 a 7796 c 7781 g 12203 t
 ORIGIN

alignment_scores:

Quality: 602.50 Length: 128
 Ratio: 4.898 Gaps: 1
 Percent Similarity: 96.094 Percent Identity: 89.062

alignment_block:

US-09-327-750D-32 x HSU80B1 ..

Align seg 1/1 to: HSU80B1 from: 1 to: 41029

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 |||||
 18412 ATGGAGTCCAAAGAGGAAACGAGCGTTAAACAATCTCATCTGGAATGT 18461
 17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
 :|||
 18462 CAACCCAGAAATGATGAAAAAGATGAAAGGAGCAAGTTGCTAATAAG 18511
 31 lyGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
 |||||
 18512 GGGAGCCCTGGCCCTACCTTTGAATTTAGTGAATACTGTGTGCCTAGA 18561
 48 GlyAsnArgArgPheProValArgGlnProIleLeuGlnTyrArgTr 64
 |||||
 18562 GGAACCGTAGGCGGTTCCCGCTTAGCGAGCCATCTCGCAGTATAGT 18611
 64 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 81
 |||||
 18612 GGACATAATCATAGCTTGAGAGCCACAGGCAAGATGAGAGAGAGA 18661
 81 snMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArg 97
 |||||
 18662 ATATGGAAGGATTTGGGAGGAGTGAGACAGCTGATGGAAGCTCAGG 18711
 98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 114
 |||||
 18712 GAAAGCAGTTGAGTCATAGTTTCGCGGCAGTCAGCACTGATCCCCCTCA 18761
 114 sHisAspHisAspGluPheCysLeuMetPro 125


```
/partial
/Note="Alu repeat: matches 298. .108 of consensus"
10593. .10688
/partial
/Note="Alu repeat: matches 96. .1 of consensus"
10698. .10781
/Note="L1 element fragment"
11112. .11403
/Note="Alu repeat: matches 1. .308 of consensus"
11922. .12004
/Note="L1 element fragment"
12062. .12308
/Note="L1 element fragment"
12423. .12707
/partial
/Note="Alu repeat: matches 302. .1 of consensus"
12710. .12916
/Note="L1 element fragment"
12979. .13081
/Note="L1 element fragment"
13473. .13532
/Note="MLT2A1 element fragment"
15871. .15928
/Note="L1 element fragment"
18180. .18417
/Note="L1 element fragment"
18778. .19056
/Note="Alu repeat: matches 1. .308 of consensus"
22279. .22350
/Note="L1 element fragment"
22449. .22995
/Note="L1 element fragment"
23239. .23374
/Note="L1 element fragment"
23371. .23606
/Note="MLT1B element fragment"
23395. .23541
/Note="MLT1A element fragment"
23705. .23778
/Note="MLT1A element fragment"
23805. .23960
/Note="L1 element fragment"
24011. .24541
/Note="L1 element fragment"
24638. .24770
/partial
/Note="Alu repeat: matches 150. .1 of consensus"
24829. .30975
/Note="L1 element fragment"
BASE COUNT 9462 a 6147 c 6129 g 9583 t
ORIGIN
```

alignment_scores:
Quality: 602.50 Length: 128
Ratio: 4.898 Gaps: 1
Percent Similarity: 96.094 Percent Identity: 89.062

alignment_block:
US-09-327-750D-32 x HSV870H8 ..

Align seg 1/1 to: HSV870H8 from: 1 to: 31321

- 1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
- 3508 ATGAGTGCCAAAGAGGAGGAGCGCTTAACAATCTCATCTGCGAAATGT 3557
|||||
- 17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
:|||||
- 3558 CAACGAGAAATGATGAAAGATGAAAGGAGCAAGTGTCTATTAAG 3607
:|||||
- 31 lyGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
|||||

```
3608 GGGAGCCCTTGGCCCTACCTTTGAATGTTAGTGAATACTGTGTGCCTAGA 3657
48 GlyAsnArgArgArgPheProValArgGlnProIleLeuGlnTyrArgTr 64
|||||
3658 GGAACCGTAGGCGGTCCCGTTAGGACGCCCATCTCGAGTATAGATG 3707
64 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 81
|||||
3708 GGACATAATGCATAGGCTTGAGAGGCCACAGGCAAGGATGAGAGAGGAGA 3757
81 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 97
|||||
3758 ATATGGAAGGATTTGGGAGAGGTGAGACAGCTGATGGAAGAGCTGAGG 3807
98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 114
|||||
3808 GAAAGCAGTTGAGTCATAGTCTGCGGGCAGTCAGCACTGATCCCCCTCA 3857
114 sHisAspHisAspGluPheCysLeuMetPro 125
|||||
3858 CCATGACCATCAGCATGAGTTTTCCTTATGCCC 3891
seq_name: gb_pr:AL133348
```

seq_documentation_block:
LOCUS AL133348 40584 bp DNA PRI 06-MAR-2000
DEFINITION Human DNA sequence from clone RPL-79P11 on chromosome Xq21.32-22.1.
Contains the gene for a novel protein similar to mouse Bex2
(brain-expressed x-linked protein 2), ESTs, STSs, GSSs and a
putative CpG island, complete sequence.
AL133348
VERSION AL133348.8 GI:7076386
KEYWORDS HTG; Bex2; CpG island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Wilson, S.
Direct Submission
Submitted (02-MAR-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Feb 24, 2000 this sequence version replaced gi:6997869.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; SW., SWISSPROT; Tr., TrEMBL; Wp., WormPEP; Information
on the WormPEP database can be found at
http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
RPL-79P11 is from the library RPCI-1 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
IMPORTANT: This sequence is not the entire insert of clone
RPL-79P11 it may be shorter because we only sequence overlapping
sections once, or longer because we arrange for a small overlap
between neighbouring submissions.

```

64 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 81
|||||
301 GGACATAATCATAGGCTGGAGAGCCACAGGCAAGGATGAGAGAGAGA 350
|||||
81 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 97
|||||
351 ATATGGAAGAGATTGGGAGAGAGGTGAGACAGCTGTATGGAAGAGCTGAG 400
|||||
98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 114
|||||
401 GAAAGCAGTTGAGTCATAGTCTCGGGCAGTCAGCACTGATCCCCCTCA 450
|||||
114 sHisAspHisHisAspGluPheCysLeuMetPro 125
|||||
451 CCATGACCATCAGCATGAGTTTGCCTTATGCCC 484
|||||

```

seq_name: gb_pat:AX078272

```

seq_documentation_block:
LOCUS AX078272 898 bp DNA PAT 22-FEB-2001
DEFINITION Sequence 76 from Patent WO0107471.
ACCESSION AX078272
VERSION AX078272.1 GI:13157963
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Hillman,J.L., Lai,P., Tang,Y.T., Yue,H., Au-Young,J., Bandman,O.,
Azimzai,Y., Yang,J., Lu,D.A., Baughn,M.R., Patterson,C. and Shah,P.
TITLE Cell cycle and proliferation proteins
JOURNAL Patent: WO 0107471-A 76 01-FEB-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..898
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="incyte ID No: 5664154CB1"
BASE COUNT 250 a 186 c 251 g 211 t
ORIGIN

```

```

alignment_scores:
Quality: 602.50 Length: 128
Ratio: 4.898 Gaps: 1
Percent Similarity: 96.094 Percent Identity: 89.062

```

alignment_block:

US-09-327-750D-32 x AX078272 ..

Align seg 1/1 to: AX078272 from: 1 to: 898

```

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
254 ATGAGTCCAAAGAGGACGAGGCTTAACAATCTCATCTGTGGAATGT 303
|||||
17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
|||||
304 CAACCCAGGAAATGATGAAAGAGATGAAAGGAGCAAGTTGCTTAATAAG 353
|||||
31 lyGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
|||||
354 GGGAGCCCTTGGCCCTACCTTGAATGTTAGTGAATCTGTGCGCTAGA 403
|||||
48 GlyAsnArgArgArgPheProLeuArgGlnProIleLeuGlnTyrArgTr 64
|||||
404 GGAACCGTAGGCGGTTCCTCCGTTAGGCAGCCCATCTCGCAGTATAGATG 453
|||||
64 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 81
|||||
454 GGACATAATCATAGGCTGGAGAGCCACAGGCAAGGATGAGAGAGAGA 503
|||||

```

```

81 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 97
|||||
504 ATATGGAAGAGATTGGGAGAGAGGTGAGACAGCTGTATGGAAGAGCTGAGG 553
|||||
98 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 114
|||||
554 GAAAGCAGTTGAGTCATAGTCTCGGGCAGTCAGCACTGATCCCCCTCA 603
|||||
114 sHisAspHisHisAspGluPheCysLeuMetPro 125
|||||
604 CCATGACCATCAGCATGAGTTTGCCTTATGCCC 637
|||||

```

seq_name: gb_pr:HSV870H8

```

seq_documentation_block:
LOCUS HSV870H8 31321 bp DNA PRI 23-NOV-1999
DEFINITION Human DNA sequence from cosmid V870H8, between markers DXS366 and
DXS87 on chromosome X contains ESTs.
ACCESSION Z70233
VERSION Z70233.1 GI:12335542
KEYWORDS X.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 31321)
AUTHORS Whiteley,M.
JOURNAL Direct Submission
Submitted (19-MAR-1995) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT
IMPORTANT: This sequence is not the entire insert of clone V870H8.
It may be shorter because we only sequence overlapping sections
once, or longer because we arrange for a small overlap between
neighbouring submissions.
The true left end of clone V870H8 is at 1 in this sequence. The
true left end of clone V693A8 is at 3121.
V870H8 is from the human chromosome X-specific cosmid library.
FEATURES
Location/Qualifiers
1..31321
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="X"
/clone="GHC-870H8"
/clone_lib="SCcv"
475..701
/note="L1 element fragment"
802..988
/note="L1 element fragment"
2533..4138
/note="L1 element fragment"
/note="match: multiple ESTs"
4189..4246
/note="29 copies of 2 mer 91 & conserved"
7496..7874
/note="L1 element fragment"
8022..8108
/note="L1 element fragment"
8302..8493
/note="L1 element fragment"
8585..8671
/note="L1 element fragment"
8769..9010
/note="L1 element fragment"
9259..9384
/note="L1 element fragment"
9624..9731
/note="MLTIC element fragment"
9781..9948
/note="MLTIC element fragment"
9890..9948
/note="MLTIB element fragment"
10418..10592

```

```

seq_name: gb_pr:AF237783
seq_documentation_block:
LOCUS AF237783 791 bp mRNA PRI 02-SEP-2000
DEFINITION Homo sapiens brain-expressed protein BEX1 (BEX1) mRNA, complete cds.
ACCESSION AF237783
VERSION AF237783.1 GI:9963898
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 791)
AUTHORS Yang, Q.S., Ying, K., Xie, Y., and Mao, Y.M.
TITLE A Novel Human X-linked Brain Expressed Protein
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 791)
AUTHORS Mao, Y.M., Xie, Y., Yang, Q.S., Wu, H., Lin, S. and Ying, K.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2000) Genetic Research Institute, Fudan University, 220 Handan Rd., Shanghai 200433, P.R.China
FEATURES
source
1..791
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="between DXS990 and DX1059 by RH (Stanford G3)"
1..791
/gene="BEX1"
172..549
/gene="BEX1"
/codon_start=1
/product="brain-expressed protein BEX1"
/protein_id="AAG05752.1"
/db_xref="GI:9963898"
/translation="MESKEKLAVNSLSMENANQENEEKEQVANKGEPLALPDAGEYC
VPRGNRRFRVROPILQYRWDMHRLGEPQARMREEMERIGEVRQLMEKLRKQLS
HSLRAVSTDPHHDHDEFLMP"
BASE COUNT 214 a 172 c 218 g 187 t
ORIGIN

alignment_scores:
Quality: 639.00 Length: 125
Ratio: 5.238 Gaps: 0
Percent Similarity: 97.600 Percent Identity: 96.800

alignment_block:
US-09-327-750D-32 x AF237783
Align seg 1/1 to: AF237783 from: 1 to: 791

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
172 ATGGAGTCCAAAGAGAACTAGCAGTAACAGCTCAGACATGCAAAATGC 221

17 aAsnGlnGluAsnGluLysGluGlnValAlaAsnLysGlyCysProL 34
|||||
222 CAACCAAGAAATTAAGAAAGGAGCAAGTTGCTTAATAAAGGGGAGCCCT 271

34 euAlaLeuProLeuAspAlaGlyTyrCysValProArgGlyAsnArg 50
|||||
272 TGGCCCTCCCTTGGATGCTGGTGAATACTGTGTGCCTAGAGGAATCGT 321

51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67
|||||
322 AGGCGGTTCGCGGTAGGCACCCCATCTGCAGTATAGATGGGATATGAT 371

67 tHisArgLeuGlyCysProGlnAlaArgMetArgGluGluAsnMetGluA 84
|||||
372 GCATAGGCTTGGAGAACACAGGCAAGGATGAGAGAGAAATATGAAA 421

```

```

84 rgileGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
|||||
422 GGATGGGGAGGGGTGAGACAGCTGATGAAAACCTGAGGGAACAGCAG 471

101 LeuSerHisSerLeuArgAlaValSerThrAspProProHisHisAspHi 117
|||||
472 TTGAGTCATAGTCTGCGGGCAGTCAGCACTGACCCCTCACCATGACCA 521

117 SHISAspGluPheCysLeuMetPro 125
|||||
522 TCATGATGAGTTTGGCTTATGCCC 546

seq_name: gb_pr:AF251053
seq_documentation_block:
LOCUS AF251053 744 bp mRNA PRI 15-APR-2001
DEFINITION Homo sapiens X-linked protein mRNA, complete cds.
ACCESSION AF251053
VERSION AF251053.1 GI:13625167
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 744)
AUTHORS Mao, Y., Xie, Y., Zhou, Z., Zhao, W., Wang, W., Huang, Y.,
Wang, S., Tang, R., Chen, X. and Wu, C.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2000) Institute of Genetics, School of Life
Sciences, Fudan University, 220 Handan Road, Shanghai 200433, P.R.
China
FEATURES
Location/Qualifiers
1..744
/organism="Homo sapiens"
/db_xref="taxon:9606"
101..487
/codon_start=1
/product="X-linked protein"
/protein_id="AAK34943.1"
/db_xref="GI:13625168"
/translation="MESKEERALNNLIVENQENDEKEQVANKGEPLALPLNVS
EYCVPRGNRRFRVROPILQYRWDMHRLGEPQARMREEMERIGEVRQLMEKLRK
QLSHSLRAVSTDPHHDHDEFLMP"
BASE COUNT 222 a 134 c 208 g 180 t
ORIGIN

alignment_scores:
Quality: 602.50 Length: 128
Ratio: 4.898 Gaps: 1
Percent Similarity: 96.094 Percent Identity: 89.062

alignment_block:
US-09-327-750D-32 x AF251053
Align seg 1/1 to: AF251053 from: 1 to: 744

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
101 ATGGAGTCCAAAGAGAAAGAGCGCTTAACAATCATCTGCTGCTAATAAG 150

17 aAsnGlnGluAsn.....GluGluLysGluGlnValAlaAsnLysG 31
|||||
151 CAACCAAGAAATGATGAAAAGATGAAAGAGCAAGTTCCTAATAAAG 200

31 lyGluProLeuAlaLeuProLeuAspAlaGlyGluTyrCysValProArg 47
|||||
201 GGGAGCCCTTGGCCCTACCTTGAATGTTAGTGAATACTGTGTGCCTAGA 250

48 GlyAsnArgArgPheProValArgGlnProIleLeuGlnTyrArgTr 64
|||||
251 GGAACCGTAGGCGGTTCCCGCTTAGCGCCATCTCGCAGTATAGATG 300

```

RPI-198P4 It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.
The true left end of clone LLOXNC01-221F2 is at 35615 in this sequence. The true right end of clone LLOXNC01-73E8 is at 100 in this sequence.

FEATURES

source

Location/Qualifiers
1..35714
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q22"
/clone="RPI-198P4"
/clone.lib="RPCI-I"

```

repeat_region      1..258
                    /note="LIP3 repeat: matches 5666. .5925 of consensus"
repeat_region      256..537
                    /note="LIP3 repeat: matches 5391. .5672 of consensus"
repeat_region      1491..1749
                    /note="AluSc repeat: matches 39. .297 of consensus"
repeat_region      2182..3810
                    /note="LIP3 repeat: matches 4533. .6150 of consensus"
repeat_region      5169..5687
                    /note="LIMA2 repeat: matches 4798. .6308 of consensus"
repeat_region      6976..7665
                    /note="LIMA2 repeat: matches 4113. .4798 of consensus"
repeat_region      8039..9213
                    /note="LIMA2 repeat: matches 2945. .4113 of consensus"
repeat_region      9523..10437
                    /note="LIMA2 repeat: matches 2016. .2945 of consensus"
repeat_region      10736..11494
                    /note="LIMA2 repeat: matches 1262. .2016 of consensus"
repeat_region      11807..12251
                    /note="LIMC repeat: matches 310. .755 of consensus"
repeat_region      12259..12331
                    /note="HERVFB21 repeat: matches 5. .77 of consensus"
repeat_region      12332..12796
                    /note="LTR21A repeat: matches 1. .490 of consensus"
repeat_region      16162..16371
                    /note="105 copies 2 mer tt 55% conserved"
repeat_region      17121..17424
                    /note="MER33 repeat: matches 1. .322 of consensus"
repeat_region      18259..18290
                    /note="16 copies 2 mer ca 87% conserved"
repeat_region      18505..18672
                    /note="MER63A repeat: matches 1. .181 of consensus"
repeat_region      19478..19814
                    /note="MER2 repeat: matches 1. .345 of consensus"
repeat_region      21061..21101
                    /note="HERVFB21 repeat: matches 42. .85 of consensus"
repeat_region      21247..21781
                    /note="MLTIE repeat: matches 44. .568 of consensus"
misc_feature       24465..24492
                    /note="Single clone region"
misc_feature       24531..24562
                    /note="Weak data"
misc_feature       24563
                    /note="Single clone region"
misc_feature       24711..24865
                    /note="Single clone region"
misc_feature       24758..25460
                    /note="region covered by subclones from a PAC DNA PCR product only at 12x coverage."
misc_feature       24875..24940
                    /note="region covered by subclones from a PAC DNA PCR product only at 12x coverage."
misc_feature       26795..26926
                    /note="CpG island"
repeat_region      24875..24940
                    /evidence="not_experimental"
repeat_region      26795..26926
                    /note="33 copies 2 mer cc 66% conserved"
repeat_region      26795..26926
                    /note="66 copies 2 mer ct 63% conserved"

```

```

repeat_region      27088..27133
                    /note="23 copies 2 mer tg 100% conserved"
repeat_region      27461..27777
                    /note="LIMB8 repeat: matches 5851. .6171 of consensus"
repeat_region      27802..28232
                    /note="LIMC/D repeat: matches 5192. .5403 of consensus"
repeat_region      28536..28930
                    /note="LIMC/D repeat: matches 4802. .5192 of consensus"
repeat_region      29240..29341
                    /note="TIGGER1 repeat: matches 484. .587 of consensus"
repeat_region      29603..30740
                    /note="TIGGER1 repeat: matches 587. .1757 of consensus"
repeat_region      31027..31365
                    /note="TIGGER1 repeat: matches 1757. .2105 of consensus"
repeat_region      31684..31691
                    /note="TIGGER1 repeat: matches 2105. .2132 of consensus"
repeat_region      31809..32015
                    /note="TIGGER1 repeat: matches 2132. .2246 of consensus"
repeat_region      32323..32389
                    /note="TIGGER1 repeat: matches 2246. .2313 of consensus"
repeat_region      34353..34636
                    /note="MLT1-INTERNAL repeat: matches 607. .889 of consensus"
repeat_region      35149..35376
                    /note="MLT1-INTERNAL repeat: matches 1001. .1233 of consensus"
repeat_region      35523..35670
                    /note="MLTIC repeat: matches 8. .159 of consensus"
BASE COUNT        9708 a 7094 c 6805 g 12107 t
ORIGIN

alignment_scores:
    Quality: 653.00      Length: 125
    Ratio: 5.266        Gaps: 0
    Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:
US-09-327-750D-32 x HS198P4/rev ..

Align seg 1/1 to reverse of: HS198P4 from: 1 to: 35714

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
24358 ATGGAGTCCAAAGAGAAACGACGAGTAAACAGTCTCAGCATGGAAATGC 24309
|||||
17 aAsnGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGlnProL 34
|||||
24308 CAACCAAGAAATGAAGAAAGGAGCAAGTTGCTAATAAGGGAGCCCT 24259
|||||
34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
|||||
24258 TGSCCCTCCCTTTGGATGCTGTGTAATCTGTGTGCTAGAGGAAATCGT 24209
|||||
51 AtgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67
|||||
24208 AGCGGTTCCGGCTTAGGCAGCCCATCTGCAGTATAGATGGGATGATG 24159
|||||
67 tHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
|||||
24158 GCATAGGCTTTGGAGAACACACAGCAGGATGAGAGAGAGAAATATCGAAA 24109
|||||
84 rGtIleGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
|||||
24108 GGATTGGGAGGAGGTGAGACAGCTGATGGAAGCTGAGGGAAGAGCAG 24059
|||||
101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 117
|||||
24058 TTGAGTCATAGTCTCGCGGAGTCAGCAGTACGCCCTCACCCTGACCA 24009
|||||
117 sHisAspGluPheCysLeuMetPro 125
|||||
24008 TCATGATGAGTTTGGCTTATGCC 23984

```

XX The present invention relates to a purified polypeptide capable of
 CC binding neurotrophin receptor (p75-NTR). The invention is useful for
 CC binding and modulating the activity of p75NTR. The peptide mediates
 CC NGF-induced apoptosis, which plays an important role in neurogenetic
 CC diseases. The peptide of the invention and p75NTR are useful for
 CC inhibiting NF-kappaB activation in a cell or a subject, for inducing
 CC caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase
 CC and fragment nuclear DNA in a cell by co-expression of (I) and p75-NTR.
 XX
 SQ Sequence 891 BP; 251 A; 182 C; 224 G; 234 T; 0 other;

alignment_scores:
 Quality: 190.00 Length: 128
 Ratio: 2.346 Gaps: 5
 Percent Similarity: 63.281 Percent Identity: 40.625

alignment_block:

us-09-327-750d-31.x AAF23529 ..

Align seg 1/1 to: AAF23529 from: 1 to: 891

2 GluSerLysGluArgAlaLeuAsnAsnLeuIleValGluAsnValas 18
 |||:||||: |||:||||:||||:||||: |||:||||: |||:||||:
 282 GAAAAACACACGAGAAAA.....AAAAATCTCATGCGCAATATTTCA 325
 18 nGlnGluAsnAspGluLysAspGluGlnValAlaAsnLysGly 35
 :|||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
 326 CCAGGAAAC.....GAAGAGATGGAGCAGCTATGCAGAAATGGAG 366
 35 luProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArgGly 51
 |||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
 367 AGGAA...GACCCCTTTGGAGAGGTGAAGGCCACCGCTGCAGGA 413
 52 AsnArgArgArgPheArgValArgGlnProIleLeuGlnTyrArgTrpas 68
 |||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
 414 AATCGAGGGGA...CAGGCTCGCCGACTTGCCTTAATTTTCGATGGGC 460
 68 pileMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnM 85
 |||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
 461 CATACCAATAGGACAGATCAATCATGGATGGGTGGAGATGGAGATCATCA 510
 85 etGluArgIleGlyGluValArgGlnLeuMetGluLysLeuArgGlu 101
 |||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
 511 TGGAAATTTTCATGGAGAGATCAGAAATCAGAGAAACTTAGGGAG 560
 102 LysGlnLeuSerHisSerLeuArgAlaValSerThrAsp...ProProHi 117
 |||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
 561 CTGCAGTTGAGGAATTTCTCGCTATCTTATGGGGAGCTCTCTAATCA 610
 117 sHisAspHisHisAspGluPheCysLeuMetPro 128
 |||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
 611 CCATGACCATCATGATGAATTTTGCCTTATGCCT 644

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517

seq_documentation_block:

ID_AAH03517 standard; cDNA; 865 BP.

XX
 AC AAH03517;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA clone (5'-primer) SEQ ID NO:352.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.

XX

PF 28-JUL-2000; 2000EP-0116126.

XX

PR 29-JUL-1999; 99JP-0248036.

PR

PR 27-AUG-1999; 99JP-0300253.

PR

PR 11-JAN-2000; 2000JP-0118776.

PR

PR 02-MAY-2000; 2000JP-0183767.

PR

PR 09-JUN-2000; 2000JP-0241899.

XX

PA (HELI-) HELIX RES INST.

XX

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX

DR WPI; 2001-318749/34.

XX

PS Claim 1; SEQ ID 352; 2537pp + CD ROM; English.

XX

CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX

SQ Sequence 865 BP; 239 A; 177 C; 224 G; 221 T; 4 other;

alignment_scores:

Quality: 187.00 Length: 129
 Ratio: 2.226 Gaps: 4
 Percent Similarity: 65.116 Percent Identity: 37.984

alignment_block:

US-09-327-750D-31 x AAH03517 ..

Align seg 1/1 to: AAH03517 from: 1 to: 865

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
 |||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
 191 ATGAGTCCAAAGAGGAACCTAGCGCAACAATCTCAACGGGGAATGCG 240
 17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
 :|||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
 241 CCAACAAGAAAC.....GAAGGAGGGGAGCAGGCCCCACGACGAGA 281
 34 lYGluproLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
 :|||:||||: |||:||||: |||:||||: |||:||||: |||:||||:
 282 ATGAAGAAGAAATCCCGCATTTGGAGGGGTGAAGGCCAGACGCTGGA 331
 51 GlyAsnArgArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
 |||:||||: |||:||||: |||:||||: |||:||||: |||:||||:

332 GGAATATACAGGGGGGGGAGTAGGACCTTGTCCCTAATTTTCGATG 381
 67 pAspIleMetHisArg...LeuGlyGluProGlnAlaArgMetArgGluG 83
 382 GGCCATACCTTAATAGGCATATTGAGCACAATGAGCGAGA.....G 422
 83 luAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99
 423 ATGATGTAGAAAGGTTGTAGGGCAGATGGAATCAAGAGAAAGACT 472
 100 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
 473 AGGACACAGCAGATGAGCAGCTATATCGCTTCCAAACTCCTGAACCT.. 520
 116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
 521GACAACCATATTGACTTTTGCCTCATACCT 550

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH13750

seq_documentation_block:

ID: AAH13750 standard; cDNA; 1229 BP.

AC AAH13750;

DT 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:10656.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126..

XX 29-JUL-1999; 93JP-0248036.

XX 27-AUG-1999; 95JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -

XX Claim 8: SEQ ID 10656; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 1229 BP; 351 A; 243 C; 296 G; 339 T; 0 other;

alignment_scores: Quality: 187.00 Length: 129

Ratio: 2.226 Gaps: 4

Percent Similarity: 65.116 Percent Identity: 37.984

alignment_block:

US-09-327-750D-31 x AAH13750 ..

Align seg 1/1 to: AAH13750 from: 1 to: 1229

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17

191 ATGGAGTCCAAAGAGAACTAGCGGCAACAATCTCAACGGGGAATGC 240

17 lAsnGlnGluAsnAspGluLysGluGlnValAlaAsnLysG 34

241 CCAACAAGAAAC.....GAAGGAGGGGAGCAGGCCCCGCCAGCA 281

34 lYcGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50

282 ATCAAGAACAATCCCGCCATTTGGAGGGGTGAAGCCCAAGACCTGGA 331

51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67

332 GGAATATACAGCGGGGGGAGTTAGCGACTTGTCCCTAATTTTCGATG 381

67 pAspIleMetHisArg...LeuGlyGluProGlnAlaArgMetArgGluG 83

382 GGCCATACCTTAATAGGCATATTGAGCACAATGAGCGAGA.....G 422

83 luAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99

423 ATGATGTAGAAAGGTTTGTAGGCCAGATGATGGAATCAAGAGAAAGACT 472

100 ArgGluLysGlnLeuSerHisSerHisArgAlaValSerThrAspProPr 116

473 AGGAAACAGCAGATGAGGCACATATATCGCTTCCAAACTCCTGAACCT.. 520

116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128

521GACAACCATATTGACTTTTGCCTCATACCT 550

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH85548

seq_documentation_block:

ID: AAC85548 standard; cDNA; 1364 BP.

XX AAC85548;

DT 04-JUN-2001 (first entry)

DE cDNA encoding CD1FF-4, Incyte ID No. 1990956CB1.

XX Human; cell differentiation; CD1FF; agonist; antagonist; epilepsy;
 KW cell proliferation; Alzheimer's disease; schizophrenic disorder;
 KW arteriosclerosis; cancer; atherosclerosis; diabetes mellitus; ss.

XX Homo sapiens.

XX Key

XX Location/Qualifiers

XX 250...612

THIS PAGE BLANK (USPTO)

OM of: US-09-327-750D-31 to: GenEmbl: * out_format : pfs
Date: Mar 11, 2002 3:34 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet.p2n.model -DEV=xlh
-O=/cgp2_1/USPTO.spool/US09327750/runat_11032002_101154_20319/app_query.fasta_1.1472
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOFCU=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MAPRIX=blsom62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09327750_@CGNL_1.8673
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-31
Query length: 128
Database: GenEmbl: *
Database sequences: 1472140
Database length: -341344837
Search time (sec): 4557.230000

score_list:

Sequence	Strd	Orig	zScore	EScore	Len	Documentation
gb_pr:AF251053	+	679.00	1230.71	2.8e-60	744	AF251053 Homo sapiens X-linked
gb_pat:AX078272	+	679.00	1229.04	3.5e-60	898	AX078272 Sequence 76 from Patent
gb_pr:HSV870H8	-	679.00	1197.46	2.0e-58	31321	Z70233 Human DNA sequence fr
gb_pr:AL133348	-	679.00	1195.16	2.7e-58	40584	AL133348 Human DNA sequence fr
gb_hcg:HSU80B1	+	679.00	1195.06	2.7e-58	41029	AL022169 Homo sapiens chromos
gb_pr:AF183416	+	606.50	1099.72	5.6e-53	642	AF183416 Homo sapiens ovarian
gb_pr:AF220189	-	606.50	1097.45	7.5e-53	828	AF220189 Homo sapiens uncharac
gb_pr:HS198P4	-	606.50	1063.99	5.5e-51	35714	AL008708 Human DNA sequence f
gb_pr:AF237783	+	592.50	1072.31	1.9e-51	791	AF237783 Homo sapiens brain-exp
gb_to:AF097439	+	502.50	908.13	2.6e-42	785	AF097439 Mus musculus brain exp
gb_to:AF097437	+	452.00	806.54	1.2e-36	2269	AF097437 Mus musculus Bexl pro
gb_to:AF097438	+	450.00	811.78	6.1e-37	835	AF097438 Mus musculus brain exp
gb_to:AF051347	+	450.00	811.74	6.1e-37	838	AF051347 Mus musculus REX-3 mR
gb_to:BC003254	+	388.00	687.56	5.1e-30	2901	BC003254 Mus musculus, Similar
gb_scs:G24641	-	339.00	613.70	6.6e-26	504	G24641 human STS WI-11354, sequ
gb_pr:HSV351F8	-	196.00	312.66	3.9e-09	45678	Z70719 Human DNA sequence fr
gb_pr:BC003190	+	190.00	337.75	1.5e-10	793	BC003190 Homo sapiens, p75NTR-a
gb_pr:HUMOGC	+	190.00	336.71	1.8e-10	891	M38188 Human unknown protein fr
gb_pr:AF187064	+	190.00	336.71	1.8e-10	891	AF187064 Homo sapiens p75NTR-as
gb_pr:AK000959	+	187.00	328.38	5.2e-10	1229	AK000959 Homo sapiens CDNA FL
gb_pat:AX100231	+	187.00	327.45	5.8e-10	1364	AX100231 Sequence 32 from Pat
gb_pr:AF187066	+	187.00	292.48	5.1e-08	69648	AL035494 Human DNA sequence f
gb_to:AF187066	+	183.50	327.00	6.2e-10	700	AF187066 Mus musculus p75NTR-a
gb_to:AF097440	-	183.50	325.23	7.7e-10	854	AF097440 Mus musculus brain exp
gb_scs:G35294	-	170.00	305.77	9.4e-09	477	G35294 human STS SHGC-37409, se
gb_to:AF187065	+	150.50	269.43	9.9e-07	519	AF187065 Rattus norvegicus p75N
gb_scs:G72708	-	140.50	253.04	8.1e-06	421	G72708 MARC 4953-4954:991939031
gb_scs:G23964	-	103.50	186.62	0.0406	372	G23964 human STS WI-15922, sequ
gb_cm:OATRICH	-	93.50	139.71	16.63	9344	Z18361 O.aries trichohyalin.2
gb_ba:SCH69	-	93.50	127.76	77.00	35824	AL079308 Streptomyces coelic
gb_pr:AB040971	+	92.50	144.57	8.92	4408	AB040971 Homo sapiens mRNA for
gb_hcg:AC085858	+	92.50	119.94	210.09	70358	AC085858 Homo sapiens clone R
gb_hcg:AF003575	+	90.50	109.80	771.39	146017	AF003575 Oriza sativa chrom
gb_cm:OCTRICHA	-	90.00	133.54	36.71	9121	Z19092 O.cuniculus trichohyal
gb_ba:AF010496	-	88.00	102.92	1.9e+03	189370	AF010496 Rhodobacter capsula
gb_to:AB041591	-	87.50	144.14	9.43	1657	AB041591 Mus musculus brain c
gb_pat:AX188346	+	87.00	145.40	8.02	1298	AX188346 Sequence 4041 from Pat
gb_pat:A64152	+	87.00	141.54	13.16	2004	A64152 Sequence 18 from Patent
gb_pat:A68338	+	87.00	141.54	13.16	2004	A68338 Sequence 18 from Patent
gb_pat:AR031270	+	87.00	141.54	13.16	2004	AR031270 Sequence 18 from pat

gb_pat:AR036392	+	87.00	141.54	13.16	2004	AR036392 Sequence 18 from p										
gb_pat:AR060824	+	87.00	141.54	13.16	2004	AR060824 Sequence 18 from p										
gb_pat:AR065979	+	87.00	141.54	13.16	2004	AR065979 Sequence 18 from p										
gb_pat:AR088933	+	87.00	141.54	13.16	2004	AR088933 Sequence 18 from p										
gb_pat:A64169	+	87.00	139.30	17.52	2576	A64169 Sequence 35 from Pat										
seq_name: gb_pr:AF251053																
seq_documentation_block:																
LOCUS	AF251053	744 bp	mRNA	PRI	15-APR-2001											
DEFINITION	Homo sapiens X-linked protein mRNA, complete cds.															
ACCESSION	AF251053															
VERSION	AF251053.1	GI:13625167														
KEYWORDS																
SOURCE	human.															
ORGANISM	Homo sapiens															
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;																
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.																
REFERENCE	1 (bases 1 to 744)															
AUTHORS	Mao,Y., Xie,Y., Zhou,Z., Zhao,W., Wang,W., Huang,Y., Wang,S., Tang,R., Chen,X. and Wu,C.															
TITLE	Direct Submission															
JOURNAL	Submitted (29-MAR-2000) Institute of Genetics, School of Life Sciences, Fudan University, 220 Handan Road, Shanghai 200433, P.R.															
FEATURES																
source	1..744	Location/Qualifiers														
/organism="Homo sapiens"																
/db_xref="taxon:9606"																
101..487																
/codon_start=1																
/product="X-linked protein"																
/protein_id="AAK34943.1"																
/db_xref="GI:13625168"																
/translation="MESKEERALLNLIVENQENDEKEQVANKGEPLALPLNVS EYCVGRNRRFRVQPILOYRWDIMHRLGEPQARMRENERIGEVRQLMEKLREK QLSHSLRAVSTDPHHDHDFCLMP"																
BASE COUNT	222 a	134 c	208 g	180 t												
ORIGIN																
alignment_scores:																
Quality:		679.00	Length:	128												
Ratio:		5.305	Gaps:	0												
Percent Similarity:		100.000	Percent Identity:	100.000												
alignment_block:																
US-09-327-750D-31 x AF251053																
Align seg 1/1 to: AF251053 from: 1 to: 744																
1	MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa	17														
101	ATGGAGTCCAAAGAGAAAGAGCGGTTTAAACAATCTCATCTGGAAAATGT	150														
17	lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG	34														
151	CRACACGAGAAATGATGAAAAAGATCAAAAGGACGAGTTGCTAATAAG	200														
34	lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg	50														
201	GGGAGCCCTTGGCCCTACCTTGAATGTTAGTGAATCTGTGTGCTAGTA	250														
51	GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr	67														
251	GGAAACCGTAGGCGGTTCCGCGTTAGGACGCCCATCTGCAGTATAGTG	300														
67	pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu	84														
301	GGACATAATGATAGGCTTGGAGAGCCAGGCAAGATGAGAGAGAGA	350														
84	snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg	100														

351 ATATTGAAAGGATTGGGAGGAGGTGAGACAGCTGATGGAAGAACTGAGG 400
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||
401 GAAAGCAGTTGAGTCATAGTCTCGGGCAGTCAGCACTGATCCCCCTCA 450
|||||

117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
451 CCATGACCATCAGCATGAGTTTGCCTTATGCCC 484
|||||

seq_name: gb_pat:AX078272

seq_documentation_block:
LOCUS AX078272 898 bp DNA PAT 22-FEB-2001
DEFINITION Sequence 76 from Patent WO0107471.

ACCESSION AX078272
VERSION AX078272.1 GI:13157963

KEYWORDS
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 898)
AUTHORS Hillman,J.L., Lai,P., Tang,Y.T., Yue,H., Au-Young,J., Bandman,O.,
Azimzai,Y., Yang,J., Lu,D.A., Baughn,M.R., Patterson,C. and Shah,P.
TITLE Cell cycle and proliferation proteins
JOURNAL Patent: WO 0107471-A 76 01-FEB-2001;

INCYTE Genomics, Inc. (US)

FEATURES
Location/Qualifiers

source
1..898
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 250 a 186 c 251 g 211 t
ORIGIN

alignment_scores:
Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-31 x AX078272 ..

Align seg 1/1 to: AX078272 from: 1 to: 898

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||

254 ATGGAGTCCAAAGAGGAACGAGCGTTAAACAATCTCATCGTGGAAAAATGT 303
|||||

17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||

304 CAACACGAGAAATGATCAAAAGATGAAAGGAGCAAGTTGCTTAATAAAG 353
|||||

34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||

354 GGGAGCCCTGGGCCCTACCTTGAATGTTAGTGAATCTGTGTGCTCTAGA 403
|||||

51 GlyAsnArgArgArgPheArgValArgGlnProIleLeuGluTyrArgTr 67
|||||

404 GGAACCGTAGGCGGTTCGCGCTTAGCGACGCCCATCTCGCAGTATTAGATG 453
|||||

67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84
|||||

454 GGACATAATGATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGAGAGA 503
|||||

84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
|||||

504 ATATGAAAGGATTGGGAGGAGTGCAGACAGCTGATGGAAGAGCTGAGG 553
|||||

101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||

554 GAAAGCAGTTGAGTCATAGTCTCGGGCAGTCAGCACTGATCCCCCTCA 603
|||||

117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||

604 CCATGACCATCAGCATGAGTTTGCCTTATGCCCC 637
|||||

seq_name: gb_pr:HSV870H8

seq_documentation_block:

LOCUS HSV870H8 31321 bp DNA PRI 23-NOV-1999
DEFINITION Human DNA sequence from cosmid V870H8, between markers DXS366 and
DXS87 on chromosome X contains ESTs.

ACCESSION Z70233
VERSION Z70233.1 GI:1235542

KEYWORDS X.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 31321)

AUTHORS Whiteley,M.

JOURNAL Direct Submission

COMMENT Submitted (19-MAR-1995) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, UK. E-mail enquires: humquery@sanger.ac.uk
IMPORTANT: This sequence is not the entire insert of clone V870H8.
It may be shorter because we only sequence overlapping sections
once, or longer because we arrange for a small overlap between
neighbouring submissions.

The true left end of clone V870H8 is at 1 in this sequence. The
true left end of clone V693A8 is at 31221.
V870H8 is from the human chromosome X-specific cosmid library.

FEATURES

Location/Qualifiers

1..31321
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="X"
/clone="GHC-870H8"
/clone_lib="SCCV"

repeat_region 475..701
/note="L1 element fragment"

repeat_region 802..988
/note="L1 element fragment"

misc_feature 2533..4138
/note="match: multiple ESTs"

repeat_region 4189..4246
/note="29 copies of 2 mer 91 % conserved"

repeat_region 7496..7874
/note="L1 element fragment"

repeat_region 8022..8108
/note="L1 element fragment"

repeat_region 8302..8493
/note="L1 element fragment"

repeat_region 8585..8671
/note="L1 element fragment"

repeat_region 8769..9010
/note="L1 element fragment"

repeat_region 9259..9384
/note="L1 element fragment"

repeat_region 9624..9731
/note="MLTIC element fragment"

repeat_region 9781..9948
/note="MLTIC element fragment"

repeat_region 9890..9948
/note="MLTIB element fragment"

repeat_region 10418..10592
/partial

repeat_region 10593..10688
/note="Alu repeat: matches 298..108 of consensus"

repeat_region 10698..10781
/note="Alu repeat: matches 96..1 of consensus"

repeat_region 10781..10871
/note="L1 element fragment"

```
repeat_region 11112..11403
/note="Alu repeat: matches 1..308 of consensus"
repeat_region 11922..12004
/note="L1 element fragment"
repeat_region 12062..12308
/note="L1 element fragment"
repeat_region 12423..12707
/partial
/note="Alu repeat: matches 302..1 of consensus"
repeat_region 12710..12916
/note="L1 element fragment"
repeat_region 12979..13081
/note="L1 element fragment"
repeat_region 13473..13532
/note="MLT2A1 element fragment"
repeat_region 15871..15928
/note="L1 element fragment"
repeat_region 18180..18417
/note="L1 element fragment"
repeat_region 18778..19056
/note="Alu repeat: matches 1..308 of consensus"
repeat_region 22279..22350
/note="L1 element fragment"
repeat_region 22449..22995
/note="L1 element fragment"
repeat_region 23239..23374
/note="L1 element fragment"
repeat_region 23371..23606
/note="MLT1B element fragment"
repeat_region 23395..23541
/note="MLT1A element fragment"
repeat_region 23705..23778
/note="MLT1A element fragment"
repeat_region 23805..23960
/note="L1 element fragment"
repeat_region 24011..24541
/note="L1 element fragment"
repeat_region 24638..24770
/partial
/note="Alu repeat: matches 150..1 of consensus"
repeat_region 24829..30975
/note="L1 element fragment"
BASE COUNT 9462 a 6147 c 6129 g 9583 t
ORIGIN

alignment_scores:
  Quality: 679.00      Length: 128
  Ratio: 5.305        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
US-09-327-750d-31 x HSV870H8

Align seg 1/1 to: HSV870H8 from: 1 to: 31321

1 MetGluSerLysGluGluArgAlaLeuAsnAsnIleValGluAsnVa 17
3508 ATGAGTCCAAAGAGGAGGAGGTTAAACAATCTCATCGTGAAAAATGT 3557
|||||
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
3558 CAACACGAGAAATGATGAAAAAGATGAAAGAGGAGCAAGTTCCTATTAAG 3607
|||||
34 lYGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
3608 GGGAGCCCTGGCCCTACCTTGAATGTTAGTGAATACTGTGTGCCTAGA 3657
|||||
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
3658 GGAACCGTAGGCGGTTCGCCGTAGGACCCCATCTCGCAGTATAGATG 3707
|||||
67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluAla 84
```

```
|||||
3708 GGACATAATGCATAGCTTGGAGAGCCACAGGCAAGGATCAGAGAGAGA 3757
|||||
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
|||||
3758 ATATGGAAGGATTGGGAGGAGGTGAGACAGCTGATGGAAAGCTGAGG 3807
|||||
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||
3808 GAAACAGCTTGAGTCATAGCTCGCGGCAGTCAGCACTGATCCCCCTCA 3857
|||||
117 shiAspHisHisAspGluPheCysLeuMetPro 128
|||||
3858 CCATGACCATCAGCATGAGTTTGCCTTATGCCC 3891

seq_name: gb_pr:AL133348

seq_documentation_block:
LOCUS AL133348 40584 bp DNA PRI 06-MAR-2000
DEFINITION Human DNA sequence from clone RPI-79p11 on chromosome Xq21.32-22.1.
Contains the gene for a novel protein similar to mouse Bex2
(brain-expressed X-linked protein 2), ESTs, STSs, GSSs and a
putative CpG island, complete sequence.
ACCESSION AL133348
VERSION AL133348.8 GI:7076386
KEYWORDS HTG; Bex2; CpG island.
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 40584)
AUTHORS Wilson,S.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Feb 24, 2000 this sequence version replaced gi:6997869.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., ENBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C.elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
RPI-79p11 is from the library RPCI-1 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
IMPORTANT: This sequence is not the entire insert of clone
RPI-79p11 It may be shorter because we only sequence overlapping
sections once, or longer because we arrange for a small overlap
between neighbouring submissions.
The true left end of clone LLOXNC01-105G4 is at 40485 in this
sequence. The true right end of clone LLOXNC01-17788 is at 100 in
this sequence.
Location/Qualifiers
1..40584
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source
```


alignment_scores:

Quality: 679.00 Length: 128
 Ratio: 5.305 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-31 x AL133348/rev ..

Align seg 1/1 to reverse of: AL133348 from: 1 to: 40584

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17

|||||
 31132 ATGGATGCCAAGAGGAGGAGCGTTAAACAATCTCATCGTGAAAATGT 31083

17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34

|||||
 31082 CAACGAGGAAATGATGAAAAAGATGAAAGGAGCAAGTTGCTAATAAAG 31033

34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50

|||||
 31032 GGGAGCCCTTGGCCCTTACCTTTGAATGTTAGTGAATACTGTGTGCCTAGA 30983

51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67

|||||
 30982 GGAACCGTAGGCGGTTCGCGTTAGCGACCCATCTGCAGTATAGATG 30933

67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84

|||||
 30932 GGACATAATGCATAGCTTGGAGAGCCACAGGCAAGGATGAGAGAGGAGA 30883

84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100

|||||
 30882 ATATGGAAGGATTTGGGAGGAGGTGAGACAGCTGATGGAAAAGCTCAGG 30833

101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117

|||||
 30832 GAAAGACAGTTGAGTCATAGTTTGGCGGCAGTCAGCACTGATCCCCCTCA 30783

117 sHisAspHisHisAspGluPheCysLeuMetPro 128

|||||
 30782 CCATGACCATCAGATGAGTTTGGCTTATGCCCC 30749

seq_name: gb_htg:HSU80B1

seq_documentation_block:

LOCUS HSU80B1 41029 bp DNA HTG 10-JUL-2001
 DEFINITION Homo sapiens chromosome X clone LLOXNC01-80B1 map q22.1-22.3, ***
 SEQUENCING IN PROGRESS ***, in ordered pieces.

ACCESSION

AL022169

VERSION AL022169.3 GI:13276704

KEYWORDS HTG; HTGS_PHASE2; HTGS_CANCELLED.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 41029)

Bird.C.

Direct Submission

Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone

requests: clonequest@sanger.ac.uk

On Mar 12, 2001 this sequence version replaced gi:12750927.

----- Genome Center

Center: Sanger Centre

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: cu80B1

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator ABI; 19% of reads

Chemistry: Dye-terminator Big Dye; 80% of reads

Consensus quality: 41009 bases at least Q40

Consensus quality: 41023 bases at least Q30

Consensus quality: 41027 bases at least Q20

Insert size: 41029; sum-of-contigs

Insert size: 51751; 0.6% error; agarose-fp

Quality coverage: 9.67x in Q20 bases; sum-of-contigs Quality

coverage: 7.67x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

FEATURES

Location/Qualifiers

source

1..41029

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="X"

/map="q22.1-22.3"

/clone="LLOXNC01-80B1"

/clone_lib="LLOXNC01"

1..41029

/note="assembly_fragment:00890"

BASE COUNT 13249 a 7796 c 7781 g 12203 t

ORIGIN

alignment_scores:

Quality: 679.00 Length: 128

Ratio: 5.305 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-31 x HSU80B1

Align seg 1/1 to: HSU80B1 from: 1 to: 41029

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17

|||||
 18412 ATGGATGCCAAGAGGAGGAGCGTTAAACAATCTCATCGTGGAATGT 18461

17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34

|||||
 18462 CAACGAGGAAATGATGAAAAAGATGAAAGGAGCAAGTTGCTAATAAAG 18511

34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50

|||||
 18512 GGGAGCCCTTGGCCCTACCTTTGAATGTTAGTGAATACTGTGTGCCTAGA 18561

51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67

|||||
 18562 GGAACCGTAGGCGGTTCGCGTTAGCGACCCATCTGCAGTATAGATG 18611

67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGluA 84

|||||
 18612 GGACATAATGCATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGAGGAGA 18661

84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100

|||||
 18662 ATATGGAAGGATTTGGGAGGAGGTGAGACAGCTGATGGAAAAGCTGAGG 18711

101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117

|||||
 18712 GAAAGCAGTTGAGTCATAGTCTCGGGGAGTCAGCACTGATCCCCCTCA 18761

117 sHisAspHisHisAspGluPheCysLeuMetPro 128

|||||
 18762 CCATGACCATCAGATGAGTTTGGCTTATGCCCC 18795

seq_name: gb_pr:AF183416

seq_documentation_block:

LOCUS AF183416 642 bp mRNA PRI 02-SEP-2000

```
DEFINITION Homo sapiens ovarian granulosa cell 13.0 kDa protein hGR74 homolog
mRNA, complete cds.
ACCESSION AF183416
VERSION AF183416.1 GI:9963770
KEYWORDS ..
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 642)
AUTHORS Peng, Y., Qian, B., Tu, Y., Xu, S., Han, Z., Fu, G. and Chen, Z.
TITLE A novel gene expressed in human adrenal gland
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 642)
AUTHORS Peng, Y., Gu, W., Huang, C., Xu, S., Han, Z., Fu, G. and Chen, Z.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-1999) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China
FEATURES
Source
1..642
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="adrenal gland"
3..380
/note="hGR74-h protein"
/codon_start=1
/evidence=not_experimental
/product="ovarian granulosa cell 13.0 kDa protein hGR74
homolog"
CDS
1..642
/protein_id="AAG09685.1"
/db_xref="GI:9963771"
/translation="MESKEKRAVNSLSMENANQNEEKEQVANKGEPLALPLDAGEYC
VPRGNRRFRVRQPILOYRWDMMHRLGEPQARMRENNMERIGEEVRLMEKLRKQLS
HSLRAVSTDPHPHDDHDFCLMP"
BASE COUNT 204 a 118 c 157 g 163 t
ORIGIN
..
alignment_scores
Quality: 606.50 Length: 128
Ratio: 4.891 Gaps: 1
Percent Similarity: 96.875 Percent Identity: 89.062
alignment_block
US-09-327-750D-31 x AF183416 ..
Align seg 1/1 to: AF183416 from: 1 to: 642
1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
3 ATGGAGTCCAAAGAGAAACGAGCAGTAAACAGTCTCAGCATGGAATGC 52
|||||
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||
53 CAACCAAGAAAAAT.....GAAGAAAGAGCAAGTTGCTTAATAAAG 93
|||||
34 lYGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
94 GGGAGCCCTTGGCCCTCCCTTTGGATGCTGGTGAATACTGTGTGCCTAGA 143
|||||
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
|||||
144 GGAATCGTAGGCGTTCCCGTTAGGCAGCCCATCTCGCATATAGATG 193
|||||
67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluAla 84
|||||
194 GGATATGATCATAGGCTTGAGCAACACAGGCAAGGATGAGAGAGAGA 243
|||||
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
|||||
244 ATATGGAAGAGGATTGGGAGGAGGTGAGACAGCTGATGGAAGAGCTGAGG 293
|||||
```

```
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHI 117
|||||
294 GAAAGAGCTTGAGTCATAGTCTCGGGCAGTCAGCACTGACCCCTCA 343
|||||
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
344 CCATGACCATCATGATGAGTTTGCCCTTATGCC 377
|||||
seq_name: gb_pr:AF220189
seq_documentation_block:
LOCUS AF220189 828 bp mRNA PRI 04-MAY-2000
DEFINITION Homo sapiens uncharacterized hypothalamus protein HBEX2 mRNA,
complete cds.
ACCESSION AF220189
VERSION AF220189.1 GI:7689028
KEYWORDS SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 828)
AUTHORS Xiao, H., Song, H., Gao, G., Ren, S., Chen, Z. and Han, Z.
TITLE A novel gene expressed in human hypothalamus
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 828)
AUTHORS Xiao, H., Song, H., Gao, G., Ren, S., Chen, Z. and Han, Z.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China
FEATURES
Source
1..828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="hypothalamus"
167..544
/codon_start=1
/product="uncharacterized hypothalamus protein HBEX2"
/protein_id="AAF67654.1"
/db_xref="GI:7689029"
/translation="MESKEKRAVNSLSMENANQNEEKEQVANKGEPLALPLDAGEYC
VPRGNRRFRVRQPILOYRWDMMHRLGEPQARMRENNMERIGEEVRLMEKLRKQLS
HSLRAVSTDPHPHDDHDFCLMP"
BASE COUNT 256 a 170 c 221 g 181 t
ORIGIN
..
alignment_scores
Quality: 606.50 Length: 128
Ratio: 4.891 Gaps: 1
Percent Similarity: 96.875 Percent Identity: 89.062
alignment_block
US-09-327-750D-31 x AF220189 ..
Align seg 1/1 to: AF220189 from: 1 to: 828
1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
167 ATGGAGTCCAAAGAGAAACGAGCAGTAAACAGTCTCAGCATGGAATGC 216
|||||
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||
217 CAACCAAGAAAAAT.....GAAGAAAGAGCAAGTTGCTTAATAAAG 257
|||||
34 lYGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
258 GGGAGCCCTTGGCCCTCCCTTTGGATGCTGGTGAATACTGTGTGCCTAGA 307
|||||
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
|||||
308 GGAATCGTAGGCGTTCCCGTTAGGCAGCCCATCTCGCATATAGATG 357
|||||
```



```
repeat_region 29603..30740
/note="TRIGGER1 repeat: matches 587..1757 of consensus"
repeat_region 31027..31365
/note="TRIGGER1 repeat: matches 1757..2105 of consensus"
repeat_region 31684..31891
/note="TRIGGER1 repeat: matches 2105..2132 of consensus"
repeat_region 31809..32015
/note="TRIGGER1 repeat: matches 2132..2246 of consensus"
repeat_region 32323..32389
/note="TRIGGER1 repeat: matches 2246..2313 of consensus"
repeat_region 34353..34636
/note="MLT1-INTERNAL repeat: matches 607..889 of consensus"
repeat_region 35149..35376
/note="MLT1-INTERNAL repeat: matches 1001..1233 of consensus"
repeat_region 35523..35670
/note="MLT1C repeat: matches 8..159 of consensus"
BASE COUNT 9708 a 7094 c 6805 g 12107 t
ORIGIN

alignment_scores:
  Quality: 606.50      Length: 128
  Ratio: 4.891         Gaps: 1
  Percent Similarity: 96.875 Percent Identity: 89.062

alignment_block:
US-09-327-750d-31 x HS198P4/rev ..
Align seg 1/1 to reverse of: HS198P4 from: 1 to: 35714
1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
24358 ATGGAGTCCAAAGAGAAACAGCAGTAAACAGTCTCAGCATGGAATGC 24309
|||||
17 lAsnGlnGluAsnAspGluLysGluGlnValAlaAsnLysG 34
|||||
24308 CAACCAAGAAAT.....GAAGAAAGGAGCAAGTTGCTAATAAG 24268
|||||
34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
24267 GGGAGCCCTTGGCCCTCCCTTGGATGCTGGTGAATACTGTGCTAG 24218
|||||
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
|||||
24217 GGAATCGTAGGCGGTTCGCGGTTCGAGCAGCCCATCTGCAGTATAGTG 24168
|||||
67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu 84
|||||
24167 GGATATGATGATGCTGGAGAACACAGGCAAGGATGAGAGAGAGA 24118
|||||
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
|||||
24117 ATATGAAAGGATTTGGGAGGAGTGCAGACGCTGATGAAAGCTGAGG 24068
|||||
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||
24067 GAAAGACGTTGAGTCATAGTCTGCGGGCAGTCAGCACTGACCCCTCA 24018
|||||
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
24017 CCATGACCATCATGATGAGTTTGCCTTATGCCC 23984
|||||

seq_name: gb_pr:AF237783
seq_documentation_block:
LOCUS AF237783 791 bp mRNA PRI 02-SEP-2000
DEFINITION Homo sapiens brain-expressed protein BEX1 (BEX1) mRNA, complete cds.
ACCESSION AF237783
VERSION AF237783.1 GI:9963898
KEYWORDS
```

```
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 791)
AUTHORS Yang,Q.S., Ying,K., Xie,Y. and Mao,Y.M.
TITLE A Novel Human X-linked Brain Expressed Protein
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 791)
AUTHORS Mao,Y.M., Xie,Y., Yang,Q.S., Wu,H., Lin,S. and Ying,K.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2000) Genetic Research Institute, Fudan
University, 220 Handan Rd., Shanghai 200433, P.R.China
FEATURES
    Location/Qualifiers
        1..791
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="X"
            /map="between DXS990 and DX1059 by RH (Stanford G3)"
            1..791
            /gene="BEX1"
            172..549
            /gene="BEX1"
            /codon_start=1
            /product="brain-expressed protein BEX1"
            /protein_id="AAG09752.1"
            /db_xref="GI:9963899"
            /translation="MESKEKLVNSLMSNANQNEKEQVANKGEPIALPLDAGEYC
            VPRNRRFRVQPIQYRNDMMHRLGEPQARMRENMERIGEGVRLMEKLRKQLS
            HSLRAVSTDPHPHDHDFCLMP"
BASE COUNT 214 a 172 c 218 g 187 t
ORIGIN

alignment_scores:
  Quality: 592.50      Length: 128
  Ratio: 4.857         Gaps: 1
  Percent Similarity: 95.312 Percent Identity: 87.500

alignment_block:
US-09-327-750D-31 x AF237783 ..
Align seg 1/1 to: AF237783 from: 1 to: 791
1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
172 ATGGAGTCCAAAGAGAACTCAGTAAACAGTCTCAGCATGGAATGC 221
|||||
17 lAsnGlnGluAsnAspGluLysGluGlnValAlaAsnLysG 34
|||||
222 CAACCAAGAAAT.....GAAGAAAGGAGCAAGTTGCTAATAAG 262
|||||
34 lyGluProLeuAlaLeuProLeuAsnValSerGluTyrCysValProArg 50
|||||
263 GGGAGCCCTTGGCCCTCCCTTGGATGCTGGTGAATACTGTGCTAG 312
|||||
51 GlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrArgTr 67
|||||
313 GGAATCGTAGGCGGTTCGCGGTTCGAGCAGCCCATCTGCAGTATAGTG 362
|||||
67 pAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluGlu 84
|||||
363 GGATATGATGATGCTGGAGAACACAGGCAAGGATGAGAGAGAGA 412
|||||
84 snMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeuArg 100
|||||
413 ATATGAAAGGATTTGGGAGGAGTGCAGACGCTGATGAAAGCTGAGG 462
|||||
101 GluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||
463 GAAAGACGTTGAGTCATAGTCTGCGGGCAGTCAGCACTGACCCCTCA 512
|||||
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
```

```

|||||
513 CCATGACCATCATGATGAGTTTGCCTTATGCC 546

seq_name: gb_ro:AF097439

seq_documentation_block:
LOCUS AF097439 785 bp mRNA ROD 13-APR-1999
DEFINITION Mus musculus brain expressed x-linked protein 2 (Bex2) mRNA,
complete cds.
ACCESSION AF097439
VERSION AF097439.1 GI:4580591
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 785)
Brown,A.L. and Kay,G.F.
Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
Hum. Mol. Genet. 8 (4), 611-619 (1999)
99172070
Erratum: [[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
2 (bases 1 to 785)
Brown,A.L. and Kay,G.F.
Direct Submission
TITLE Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
JOURNAL Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
MEDLINE Location/Qualifiers
FEATURES
source
1..785
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="dbEST:W48832"
/chromosome="X"
/map="near Plp"
/dev_stage="embryo; 15.5 dpc"
1..785
/gene="Bex2"
139..528
/gene="Bex2"
/codon_start=1
/product="brain expressed x-linked protein 2"
/protein_id="AAD24430.1"
/db_xref="GI:4580592"
/translation="MESKVEQGVKNLNMENDHQKEEKQPDASKRDPVALPFEA
GDYVPRGRRRFRVQPIVIRWDLMRHVGEPQGRMRREENVQRFQDDVROLMEKLR
RQLSRSLRAVSTDPHHDHDECLMP"
BASE COUNT 212 a 163 c 235 g 175 t
ORIGIN

alignment_scores:
Quality: 502.50 Length: 129
Ratio: 4.153 Gaps: 1
Percent Similarity: 93.798 Percent Identity: 71.318

alignment_block:
US-09-327-750d-31 x AF097439 ..
Align seg 1/1 to: AF097439 from: 1 to: 785
1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||
139 ATGGAGTCCAAATGGCAACAGCGGTGAAATAATCTCAACATGAGATGA 188
|||||
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||
189 CCATCAGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 238
|||||
34 lyGluProLeu...AlaLeuProLeuAsnValSerGluTyrCysValPro 49
|||||

```

```

239 GGGATCCGATTTGGCCCTTCCAGCTGGAGACTACTACGTGCCT 288
50 ArgGlyAsnArgArgArgPheArgValArgGlnProIleLeuGlnTyrAr 66
|||||
289 AGAGGAGTGGCAGCGGTTCGGGTTCGCAGCCCATCGTGCACTACAG 338
66 gTTPAspIleMethIserArgLeuGluProGlnAlaArgMetArgGluG 83
|||||
339 ATGGGACCTGATCATAGGTTGGGAGGCCGCCAGGGAAGCATGAGAGG 388
83 luAsnMetGluArgIleGlyGluValArgGlnLeuMetGluLysLeu 99
|||||
389 AGAACGTACAGAGTTTGGGGATGATGTGACAGCTCATGTGAGAAAGCTG 438
100 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||
439 AGGAAAGGAGCTGAGCACAGCTCGCGCGGTAGCACTGACCGGCC 488
116 oHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
489 TCATCATGACCAACCATGATGAGTTTGCCTTATGCC 525
seq_name: gb_ro:AF097437

seq_documentation_block:
LOCUS AF097437 2269 bp DNA ROD 05-AUG-1999
DEFINITION Mus musculus Bex1 protein (Bex1) gene, complete cds.
ACCESSION AF097437
VERSION AF097437.1 GI:5702155
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2269)
Brown,A.L. and Kay,G.F.
Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome.
Hum. Mol. Genet. 8 (4), 611-619 (1999)
99172070
PUBMED 10072429
REFERENCE
2 (bases 1 to 2269)
Brown,A.L. and Kay,G.F.
Direct Submission
TITLE Submitted (07-OCT-1998) Cancer Unit, Queensland Institute of
JOURNAL Medical Research, Herston Rd, Brisbane, Qld 4030, Australia
FEATURES
source
1..2269
/organism="Mus musculus"
/strain="129SVJ"
/db_xref="taxon:10090"
/chromosome="X"
/map="near Plp"
/tissue_type="liver"
/join(409..509,866..942,1263..1873)
/gene="Bex1"
/product="Bex1 protein"
409..1873
/gene="Bex1"
/note="expressed in brain; X-linked"
1268..1654
/gene="Bex1"
/codon_start=1
/product="Bex1 protein"
/protein_id="AAD47168.1"
/db_xref="GI:5702156"
/translation="MESKDGQGVKNLNMENDHQKEEKQPDITIRREPVALTSEAG
KNCAPRGRRRFRVQPIAHYRWDLMRVGPQGRMRREENVQRFQDDVROLMEKLRER
QLSHSLRAVSTDPHHDHDECLMP"
BASE COUNT 563 a 554 c 705 g 447 t
ORIGIN

```


TITLE Characterization of genes which exhibit reduced expression during the retinoic acid-induced differentiation of F9 teratocarcinoma cells: involvement of cyclin D3 in RA-mediated growth arrest

JOURNAL Mol. Cell. Endocrinol. 143 (1-2), 155-166 (1998)

MEDLINE 99021197

REFERENCE 2 (bases 1 to 838)

AUTHORS Faria,T.N., LaRosa,G., Wilen,E., Liao,L. and Gudas,L.J.

TITLE Direct Submission

JOURNAL Submitted (26-FEB-1998) Pharmacology, Cornell University Medical College, 1300 York Avenue, New York, NY 10021, USA

FEATURES

source

1..838

/organism="Mus musculus"

/db_xref="taxon:10090"

/cell_line="F9 teratocarcinoma"

229..579

/note="retinoic acid reduced gene expression"

/codon_start=1

/product="REX-3"

/protein_id="AAC61929.1"

/db_xref="GI:3510643"

/translation="MENDHOKKEEKEKPODTIRREPAVALISEAGKNCAPRGRRRRFVRQPIAHYRWDLQVRGEPQGRMRNEVQFGDVRQLMKLRQLSHSLRAVSTDPPHDDHDEFLMP"

BASE COUNT 238 a 181 c 251 g 168 t

ORIGIN

alignment_scores

Quality: 450.00 Length: 129

Ratio: 3.879 Gaps: 2

Percent Similarity: 89.922 Percent Identity: 66.667

alignment_block:

US-09-327-750D-31 x AF051347

Align seg 1/1 to: AF051347 from: 1 to: 838

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17

193 ATGGAGTCCAAAGATCAA...GGCGTGAAAAATCTCAACATGGAGATGA 239

17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34

240 CCATCAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 289

34 lyGluPro...LeuAlaLeuProLeuAsnValSerGluTyrCysValPro 49

290 GGAGGCCAGCTGTGGCCCTGATCTCCGAGGCTGGCAAAACCTGTGGCCT 339

50 ArgGlyAsnArgArgPheArgValArgGlnProIleLeuGlnTyrAr 66

340 AGAGGAGTCCGAGGCGGTTCGGGTTCGGCAGCCCATCTCCTCATATAG 389

66 gTrpAspIleMetHisArgLeuGlyGluProGlnAlaArgMetArgGluG 83

390 ATGGGACCTCATCAGAGGTTGGGAGGCCAGGAGGAGGAGGAGGAGG 439

83 lAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99

440 AGAACCTACAGAGGTTGGGAGGCTGTGAGAGAGCTGATGAGAGAGCTG 489

100 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116

490 AGGGAAGGAGGAGCTGAGCCACGCTGCGGCGGTGAGCACTGACCCGCC 539

116 oHisHisAspHisAspGluPheCysLeuMetPro 128

540 TCATCATGACCACCATGATGAGTTTGCCTCATGCC 576

seq_name: gb_ro:BC003254

seq_documentation_block:

LOCUS BC003254 2901 bp mRNA ROD 20-FEB-2001

DEFINITION Mus musculus, Similar to dentatorubral pallidolysian atrophy, clone MGC:5758, mRNA, complete cds.

ACCESSION BC003254

VERSION BC003254.1 GI:13096906

KEYWORDS MGC.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2901)

TITLE Direct Submission

JOURNAL Submitted (20-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalon@bcm.tmc.edu.
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 9 Row: 0 Column: 6.

FEATURES

Location/Qualifiers

1..2901

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="MGC:5758 IMAGE:3500522"

/tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."

/clone.lib="NCI CGAP_Mam6"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

755..2359

/codon_start=1

/product="Similar to dentatorubral pallidolysian atrophy"

/protein_id="AAH03254.1"

/db_xref="GI:13096907"

/translation="MVPQFLLPGLKPPHVTPLHPRAKEHPTPSHQSLQSP PPQLPFLSPAPATTGPPLTATQIKQEPAEYEPPEPPSPKSPSPKPVVDVPS HASQSAFNKHLDRGNSCARSDLYFVPLEGSKAKRADLVEKYRREAEQRAREEKE REREREREKERERELERSVLAQEGRAPVECPGLGVPVHRPFPFGSAVATVPP YLGDPTPALRTLSEYARPHVMSPGNRNHPFYVPLGAVDPLGLLYNVPALYSDDPAARE REREARERDLORLKPGEVKPSELEPLHGVPGCLDPEPHGGGLAQPGPLHPFP FHPSLGPLERLERLALAAPALPARNYSYARERLAARQHAERVAALGNDRLQMLNVT PHHQHSIHSHSLHLHQDDAIIHAASASVHPLDIPASGSHLTRIDPYTPAGTLPNLLPH PLHENEVLRHOLFAPYRDLFPASIPASPMASAAHQAOLQAMHAQSAELQRLALEQQOQLJHAH HPLHSVLPQAQEDYYSHLKESDKPL"

BASE COUNT 623 a 958 c 791 g 529 t

ORIGIN

alignment_scores

Quality: 388.00 Length: 119

Ratio: 3.660 Gaps: 2

Percent Similarity: 89.076 Percent Identity: 63.866

alignment_block:

US-09-327-750D-31 x BC003254

Align seg 1/1 to: BC003254 from: 1 to: 2901

```

1 MetGluSerLysGluGluArgAlaLeuAsnAsnLeuIleValGluAsnVa 17
|||||.....:|||||:|||||
223 ATGGAGTCCAAAGATCAA...GGCGTGAATAATCTCAACATGGAGAATGA 269
|||||.....:|||||:|||||
17 lAsnGlnGluAsnAspGluLysAspGluLysGluGlnValAlaAsnLysG 34
|||||.....:|||||:|||||
270 CCATCAGAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 319
|||||.....:|||||:|||||
34 lyGluPro...LeuAlaLeuProLeuAsnValSerGluTyrCysValPro 49
|||||.....:|||||:|||||
320 GGGAGCCAGCTGTGGCCCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCT 369
|||||.....:|||||:|||||
50 ArgGlyAsnArgArgArgPheArgValArgGlnProIleLeuGlnTyrAr 66
|||||.....:|||||:|||||
370 AGAGGAGGTCGACAGCGGTTCCGGGTTCCGGCAGCCCTCCTACTATAG 419
|||||.....:|||||:|||||
66 gTrpAspIleMetHisArgLeuGluGluProGlnAlaArgMetArgGluG 83
|||||.....:|||||:|||||
420 ATGGGACCTGATCGACAGAGGTTGGGAGGCCCCAGGAGGATGAGAGAGG 469
|||||.....:|||||:|||||
83 luAsnMetGluArgIleGlyGluGluValArgGlnLeuMetGluLysLeu 99
|||||.....:|||||:|||||
470 AGAACGTACAGAGGTTTGGGGGTGATGTGACAGAGCTCATGGAGAAAGCTG 519
|||||.....:|||||:|||||
100 ArgGluLysGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||.....:|||||:|||||
520 AGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 569
|||||.....:|||||:|||||
116 oHisHis 118
|||||
570 TCATCAT 576

```

seq_name: gb_sts:G24641

```

seq_documentation_block:
LOCUS      G24641          504 bp    DNA          STS          31-MAY-1996
DEFINITION human STS WI-11354, sequence tagged site.
ACCESSION  G24641
VERSION    G24641.1  GI:1344967
KEYWORDS   STS; STS sequence; primer; sequence tagged site.
SOURCE     human STSs derived from sequences in dbEST and the Unigene
           collection.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1  (bases 1 to 504)
AUTHORS   Hudson,T.
TITLE     Whitehead Institute/MIT Center for Genome Research; Physically
JOURNAL   Mapped STS
COMMENT   Unpublished (1995)

```

Contact: Thomas Hudson
 Whitehead Institute/MIT Center for Genome Research
 Whitehead Institute for Biomedical Research
 9 Cambridge Center, Cambridge MA 02142 USA
 Tel: 617 252 1900
 Fax: 617 252 1902
 Email: thudson@genome.wi.mit.edu

Primer A: TTTTGTGCTTTTAAAGGTGC
 Primer B: TTACGTGGGTCTCCTATTACCA
 STS size: 150
 PCR Profile:

Presoak:
 Denaturation:
 Annealing: 56 degrees C
 Polymerization:
 PCR Cycles: 35
 Thermal Cycler:

Protocol:
 Template: 10 ng

Primer: each 5 pM
 dNTPs: each 4 nM
 Taq Polymerase: 0.025 units/ul
 Total Vol: 20 ul

Buffer:

MgCl2: 1.5 mM
 KCl: 50 mM
 Tris-HCL: 10 mM
 pH: 9.3

Derived from dbEST (genbank accession T86927).

```

FEATURES
    source              1..504
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
    STS                 1..150
    primer_bind         1..22
    primer_bind         complement(129..150)
    BASE COUNT         134 a 118 c 101 g 144 t 7 others
    ORIGIN

```

alignment_scores:
 Quality: 339.00 Length: 85
 Ratio: 4.644 Gaps: 3
 Percent Simlarity: 85.882 Percent Identity: 83.529

alignment_block:

US-09-327-750D-31 x G24641/rev ..

Align seg 1/1 to reverse of: G24641 from: 1 to: 504

```

47 CysValProArgGly.AsnArgArgArgPhe.ArgValArgGlnProIle 62
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
503 TGTCCCTTANAGGAANCCGTAGCGGTTCCCTCGTTAGGCANCCCCAC 454
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63 LeuGlnTyrArgTrpAspIleMetHisArgLeuGlyGlu.ProGlnAla 79
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
453 CGCAGTANAGATGGGATATGATGTCATAGGCTTGGAGAACCCACAGGCAA 404
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
79 rgMetArgGluGluAsnMetGluArgIleGlyGluValArgGlnLeu 95
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
403 GGATGAGAGAGAGAGATATGGAAGGATTTGGGAGGAGGTGAGACAGCTG 354
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
96 MetGluLysLeuArgGluLysGlnLeuSerHisSerLeuArgAlaValSe 112
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
353 ATGGAAAAGCTGAGGGAAGAGAGTTGAGTCANAGTCTCGGGCAGCTCAG 304
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
112 rThrAspProProHisHisAspHisHisAspGluPheCysLeuMetPro 128
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
303 CACTGACCCCCCTCACCATGACCATCATGATGAGTTTGTCTTTATGCC 255
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


ethanol-precipitation. The cDNAs were ligated to Lone-linker L1-Sal3 (include SalI sequence). The cDNAs were purified by phenol/chloroform and separated from free linkers by Centricon 100. Then, cDNAs were amplified by long-range high fidelity PCR using Takara's Ex Taq polymerase. Then, the cDNAs were purified by phenol/chloroform and by Centricon 100. The cDNAs were digested with SalI and NotI enzymes. Then, the cDNAs were size selected by Gibco's Size Fractionation Column. The cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by chemical method. The library was constructed by Xiaohong Wang and Minoru S. H. Ko. "

BASE COUNT 145 a 182 c 133 g 182 t
ORIGIN

alignment_scores:
Quality: 683.00 Length: 128
Ratio: 5.378 Gaps: 0
Percent Similarity: 99.219 Percent Identity: 99.219

alignment_block:
US-09-327-750D-30 x AW536404/rev ..

Align seg 1/1 to reverse of: AW536404 from: 1 to: 642

```

1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspH1 17
|||||
606 ATGGAGTCCAAAGATCAAGGCGTGAATAATCTCAACATGGAGAAATGACCA 557
|||||
17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleArgArgG 34
|||||
556 TCAGAAAAGAGGAGGAGAGAAAACCAAGATACCATCAGAGGG 507
|||||
34 luProLaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
|||||
506 AGCCAGCTGTGGCCCTGACCTCGGAGGCTGGCAAAACCTGTCACCTAGA 457
|||||
51 GlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67
|||||
456 GGAGGTTCGAGGCGGTTCGGGTTCCGGAGCCATCCCTCACTATAGATG 407
|||||
67 pAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluGlu 84
|||||
406 GGACCTGATGCAGAGGTTGGGGAGCCCGAGGAGGATGAGAGAGAGA 357
|||||
84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
|||||
356 ACGTACAGAGGTTTGGGGGTGATGTGAGACAGCTCATGGAGAGCTGAGG 307
|||||
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH1 117
|||||
306 GAAAGCAGCTGAGCCACAGCCTCGCGGGGTTAGCACTGACCCGCTCA 257
|||||
117 sHisAspHisHisAspPheCysLeuMetPro 128
|||||
256 TCATGACCACCATGATGAGTTTGGCTCATGCCCC 223
|||||

```

seq_name: gb_est2:BG064920

seq_documentation_block:
LOCUS BG064920 707 bp mRNA EST 26-JAN-2001
DEFINITION H3025001-3 NIA Mouse 15K cDNA Clone Set Mus musculus CDNA Clone
ACCESSION BG064920
VERSION BG064920.1 GI:12547483
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 707)

```

34 luProLaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
|||||
242 AGCCAGCTGTGGCCCTGACCTCCGAGGCTGGCAAAACCTGTCGCGCTAGA 291
|||||
51 GlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67
|||||
292 GGAGGTTCGAGGCGGTTCGGGTTCCGGAGCCATCCCTCACTATAGATG 341
|||||
67 pAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluGlu 84
|||||
342 GGACCTGATGCAGAGGTTGGGGAGCCCGAGGAGGATGAGAGAGAGA 391
|||||
84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
|||||
392 ACGTACAGAGGTTTGGGGGTGATGTGAGACAGCTCATGGAGAGCTGAGG 441
|||||
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH1 117
|||||
442 GAAAGCAGCTGAGCCACAGCCTCGCGGGGTTAGCACTGACCCGCTCA 491
|||||
117 sHisAspHisHisAspPheCysLeuMetPro 128
|||||
492 TCATGACCACCATGATGAGTTTGGCTCATGCCCC 525
|||||

```

seq_name: gb_est1:AW536404

```

seq_documentation_block: 642 bp mRNA EST 31-AUG-2000
LOCUS AW536404
DEFINITION G0104A11-3 NIA Mouse E7.5 Embryonic Portion cDNA Library Mus
musculus cDNA clone G0104A11 3', mRNA sequence.
ACCESSION AW536404
VERSION AW536404.1 GI:7178821
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 642)
AUTHORS Tanaka, T.S., Jaradat, S.A., Lim, M.K., Kargul, G.J., Wang, X., Grahovac
M.J., Pantano, S., Sano, Y., Piao, Y., Negaraja, R., Doi, H., Wood, W.H.,
III, Becker, K.G. and Ko, M.S.H.
Genome-wide expression profiling of mid-gestation placenta and
embryo using a 15,000 mouse developmental cDNA microarray
Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
20381348
CONTACT: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: G0104 row: A column: 11
Seq primer: -21M13 Forward
High quality sequence stop: 642
POLYA=Yes.

```

FEATURES
source
1..642
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G0104A11"
/clone_lib="NIA Mouse E7.5 Embryonic Portion cDNA Library"
/sex="unknown"
/dev_stage="7.5dpc Embryo"
/lab_host="DH10B"
/note="vector: pSPORT1 (Gibco/BRL Life Technology);
Site_1: SalI; Site_2: NotI; Total RNAs were extracted from
6 Embryo. The double-stranded cDNA was synthesized by
Gibco's kit with an Oligo(dT) primer (NotI primer-adaptor
from Gibco/BRL)
[5'-pGACTAGTCTAGATCCGAGCGGCCCTTTTCTTTTCTTTT-3']
from 0.51ug of mRNA. The double-stranded cDNAs were
treated with T4 DNA polymerase and purified by

AUTHORS

Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
T.S., Carter,M.G. and KO,M.S.H.
Verification and Initial annotation of NIA mouse 15K cDNA clone set
Unpublished (2001)
Other_ESTs: H3025D01-5
Contact: George J. Kargul
Laboratory of Genetics

TITLE

JOURNAL

COMMENT

National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.
Plate: H3025 row: D column: 01
Seq primer: -21M13 Forward
High quality sequence stop: 707
POLYA=Yes.

FEATURES

source

1..707
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="H3025D01"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"

/note="Vector: pSPORT1; Site_1: Salt; Site_2: NotI; This
clone is among a rearranged set of 15,247 clones from 11
embryo cDNA libraries (including preimplantation stage
embryos from unfertilized egg to blastocyst, embryonic
part of E7.5 embryos, extraembryonic part of E7.5 embryos
, and E12.5 female mesonephros/gonad) and one newborn
ovary cDNA library. Average insert size 1.5 kb. All
source libraries are cloned unidirectionally with Oligo(dT
)-Not primers. References include: (1) Genome-wide
expression profiling of mid-gestation placenta and embryo
using a 15,000 mouse developmental cDNA microarray, 2000,
Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)
Large-scale cDNA analysis reveals phased gene expression
patterns during preimplantation mouse development, 2000,
Development, 127: 1737-1749; (3) Genome-wide mapping of
unselected transcripts from extraembryonic tissue of
7.5-day mouse embryos reveals enrichment in the t-complex
and under-representation on the X chromosome, 1998, Hum
Mol Genet 7: 1967-1978."

BASE COUNT 154 a 209 c 146 g 198 t

ORIGIN

alignment_scores:

Quality: 680.00 Length: 128
Ratio: 5.354 Gaps: 0
Percent Similarity: 99.219 Percent Identity: 98.438

alignment_block:

US-09-327-750d-30 x BG064920/rev ..

Align seg 1/1 to reverse of: BG064920 from: 1 to: 707

1 luProAlaValAlaLeuLeSerGluAlaGlyLysAsnMetGluAsnAspHi 17
|||||
606 ATGGAGTCCAAAGATCAAGCGCTGAAATCTCAACATGGAGAATGACCA 557
|||||
17 sGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgArg 34
|||||
556 TCAGAAAAAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 507
|||||
34 luProAlaValAlaLeuLeSerGluAlaGlyLysAsnMetGluAsnAspHi 50
|||||
506 ACCGAGCTGTGCGCTGACCTCCGAGGCTGGCAAAACATGCGACCTAGA 457
|||||
51 GlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67

|||||
456 GGAGGTGCGAGCGGTTCGGGTTCGGCAGCCCATCGCTCACTATAGATG 407
|||||
67 pAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluGlu 84
|||||
406 GGACCTGATGCAGAGGTTGGGGAGCCCGGAGGAGGATGAGAGAGAGA 357
|||||
84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
|||||
356 ACGTACAGAGGTTGGGGGTGATGTGAGACAGCTCATGCGAGAAGCTGAGG 307
|||||
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPro 117
|||||
306 GAAAGGCACTGAGCCACAGCGTCGGGCGGTAGCACTGACCCGCTCA 257
|||||
117 shiAspHisAspGluPheCysLeuMetPro 128
|||||
256 TCATGACCACCATGATGAGTTTTCCTCATGCC 223

seq_name: gb_est2:BG277659

seq_documentation_block:

LOCUS BG277659 575 bp mRNA EST 21-FEB-2001
DEFINITION ux47c11.y1 Soares_NMAX_maxillary_process Mus musculus cDNA clone
IMAGE:3513237 5' similar to TR:Q9RLJ2 Q9RLJ2 BEX1 PROTEIN. ;, mRNA
sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

US-09-327-750D-30 x BG277659

Align seg 1/1 to: BG277659 from: 1 to: 575

```
1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspHi 17
|||||
192 ATGGAGTCCAAAGATCAAGCGCTGAAATAATCTCAACATGGAGAATGACCA 241
|||||
17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleArgArg 34
|||||
242 TCAGAAAAGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 291
|||||
34 luProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
|||||
292 AGCCAGCTGTGGCCCTGACCTCCGAGGCTGGCAAAATCTGCACCTAGA 341
|||||
51 GlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67
|||||
342 GGAGGTCCGAGGCGGTTCGGGGTTCGGCAGCCATCGCTCACTATAGATG 391
|||||
67 pAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluAla 84
|||||
392 GGACCTGTATCGAGAGGTTGGGAGGCCCCAGGAGGAGGAGGAGGAGG 441
|||||
84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
|||||
442 ACCTACAGAGTTTNGGGTGTATGTGAGACAGCTCATGGAGAGCTGAGG 491
|||||
101 GluArgGlnLeuSerHisSerHisLeuArgAlaValSerThrAspProHi 117
|||||
492 GAAAGCAGCTGAGCCACAGCTCGGGCGGTAGCACCTGAGCCCGCTCA 541
|||||
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
542 TCATGACCAACCATGATGAGTTGTGCCTCATGCC 575
|||||
```

seq_name: gb_est2:BG870503

```
seq_documentation_block: 756 bp mRNA EST 29-MAY-2001
LOCUS BG870503
DEFINITION 602791422f1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4922644 5',
mRNA sequence.
ACCESSION BG870503
VERSION BG870503.1 GI:14221043
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
NTH-MSC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10842 row: 1 column: 05
High quality sequence stop: 756.
Location/Qualifiers
1..756
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site_1:
```

FEATURES

source

Noti; Site_2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 207 a 164 c 230 g 155 t
ORIGIN

alignment_scores:
Quality: 670.00 Length: 129
Ratio: 5.276 Gaps: 1
Percent Similarity: 98.450 Percent Identity: 98.450

alignment_block:

US-09-327-750D-30 x BG870503

Align seg 1/1 to: BG870503 from: 1 to: 756

```
1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspHi 17
|||||
165 ATGGAGTCCAAAGATCAAGCGCTGAAATAATCTCAACATGGAGAATGACCA 214
|||||
17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleArgArg 34
|||||
215 TCAGAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 264
|||||
34 luProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
|||||
265 ACCCAGCTGTGGCCCTGACCTCCGAGGCTGGCAAAATCTGCACCTAGA 314
|||||
51 GlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67
|||||
315 GCAGGTCCGAGCGGTTCGGGTTCGGCAGCCATCGCTCACTATAGATG 364
|||||
67 pAspLeuMetGlnArgValGlyGluProGlnGlyArgMet.ArgGluGlu 83
|||||
365 GGACCTGTATCGAGAGGTTGGGAGGCCCCAGGAGGAGGATGACGAGGAG 414
|||||
84 AsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
|||||
415 AACGTACAGAGGTTTGGGGGTGATGTGAGACAGCTCATGGAGAGCTGAG 464
|||||
100 gGluArgGlnLeuSerHisSerHisLeuArgAlaValSerThrAspProHi 117
|||||
465 GGAAGGCGAGCTGAGCCACAGCTCGGGCGGTAGCACCTGAGCCCGCTC 514
|||||
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
515 ATCATGACCAACCATGATGAGTTTGTGCCTCATGCC 549
|||||
```

seq_name: gb_est1:BE654459

```
seq_documentation_block: 499 bp mRNA EST 06-SEP-2000
LOCUS BE654459
DEFINITION UI-M-AJ1-aha-f-10-0-UI.r1 NIH_BMAP_MOB_N Mus musculus cDNA clone
UI-M-AJ1-aha-f-10-0-UI 5', mRNA sequence.
ACCESSION BE654459
VERSION BE654459.1 GI:9980372
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 499)
AUTHORS Bonaudo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704477
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
```


Align seq 1/1 to reverse of: AW536974 from: 1 to: 577

11 LeuAsnMetGluAsnAspHisGlnIleLysGluGluLysGluGluLysPr 27
576 CTCACATCGGAGATGACCATTACAGAAAAGGAGAGAGAGAAAAGCC 527
27 oGlnAspThrIleArgArgGluProAlaValAlaLeuIleSerGluAlaG 44
526 ACNAGNATACCATCAGAGGGGAGCCAGCTGTGGCCCTGACCTCCGAGCGTG 477
44 lYlYsAsnCysAlaProArgGlyGlyArgArgPheArgValArgGln 60
476 GCAAAACCTGTGCACCTAGAGGAGGTGCAGGCGGTTCCGGTGTGGCAG 427
61 ProIleAlaHisTyrArgTriaAspLeuMetClnaArgValGlyClnuProG 77
426 CCCATCGCTCATATAGATGGAGCTCATGCAGAGGTTGGGAGGCCCA 377
77 nGlyArgMetArgGluGluAsnValGlnArgPheGlyGlyAspValArgG 94
376 GGGAGGATGAGAGAGGAGAACGTACAGAGTTTGGGGTGATGTCAGAC 327
94 lnLeuMetGluLysLeuArgGluArgGlnLeuSerHisSerLeuArgAla 110
326 AGCTCATGGGAAGTGAAGGAAAGGAGCGTGCAGCCACAGCTCGCGGGG 277
111 ValSerThrAspProHisHisAspHisHisAspClnuPheCysLeuMe 127
276 GTTAGCATGACCCGCCCTCATATGACCAACCATGATGAGTTTGGCTCAT 227
127 tPro 128
226 GCCC 223

seq_name: qb_est2:BF607762

seq_documentation_block:	785 bp	mRNA	EST	01-APR-2001
LOCUS	BF607762	Mouse 9-day fetus	CDNA library ICRFP522	Mus musculus
DEFINITION	MY1_000704	CDNA clone ICRFP522A1947 5', mRNA sequence.		
ACCESSION	BF607762			
VERSION	BF607762.1	GI:13504254		
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Yahyawi, M., Hennig, S., Neidhardt, L., Radelof, U., Hermann, B. G., Lehrach, H. and O'Brien, J.			
TITLE	Detection of a high number of novel genes in a 9-day mouse embryo CDNA library normalised by oligonucleotide fingerprinting			
JOURNAL	Unpublished (2001)			

CONTACT: Hennig S.
Laboratory 123, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Thnstr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennig@molgen.mpg.de
EST's are made from clones being representatives of clone clusters.
Clone clusters were calculated from oligonucleotide fingerprints.

```

PCR primers
  FORWARD: 5'-GAGCTATTCCAGAAGTAGTGA-3'
  BACKWARD: 5'-TAATAGACTCACTATAGGG-3'
  Seq primer: 5'-ATTAGTGACACTATAG-3'
  High quality sequence stop: 785.

FEATURES
  source
    .
    1..785
    /organism="Mus musculus"
    /db_xref="taxon:10090"
    /clone="ICRFp5242A1947"
    /clone_lib="Mouse 9-day fet

```

```

/tissue_type="whole embryo"
/dev_stage="embryonic 9-day"
/lab_host="E.coli, xli blue"
/notes="vector: pSVSPORT1; Site_1: NotI; Site_2: SalI;
Library preparation by oligo.dT priming of RNA. Clones can
be ordered from the Resource Center in Berlin,
http://www.rzod.de"

```

BASE COUNT	190 a	188 c	233 g	153 t	21 others
ORIGIN					

alignment_scores:			
Quality:	591.00	Length:	128
Ratio:	4.844	Gaps:	4
Percent Similarity:	95.312	Percent Identity:	93.750

alignment_block:

US-09-327-750D-30 x BF607762

Align seq 1/1 to: BF607762 from: 1 to: 785

1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspHis 17
259 ATGGAGTCCAAAGATCAAGCGTCAAAATCTCAACATGGAGAATGACCA 308
17 sGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgArgG 34
309 TCAGAAAAAGGAGGAGAGAGAGAAAGAACCCNCAAGATACCATCAGAAAGG 358
34 luProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
359 AGCCAGCTGTGGCCCTGACCTCGAGGCTGGCAAAACTGTGCACCTAGA 408
51 GlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67
409 GGAGGTCCGAGCGGTTCCGGGTTCCGACGCCCATCGCTCACTATAGATG 458
67 pAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluGluA 84
459 GGACCTGATGCAGAGGTTGGGAGGCCCCAGGGAAGGATGAGAGAGGAGA 508
84 snValGlnArgPheGlyGlyAspValArgGlnLeuMet.GluLysLeuAr 100
509 AGCTACAGAGTTTGGGGGTGATGTGAGACAGGTCTATCGGGGAAGCTGAG 558
100 gGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAsp.ProPro 116
559 GGAAAGGCACCTGAACCAACAGCCCTCGGGGGGTTAGCACTGACCCCGCT 608
117 His.HisAspHisHis.AspGluPhe 124
609 CATTCATGACCACCATGGATGAGTTT 634

seq name: qb est2:BG228077

seq. documentation block:			
LOCUS	BG228077	587 bp	mRNA EST
DEFINITION	ux47c11.x1 Soares_NNMAX_maxillary_process Mus musculus cDNA clone IMAGE:3513237 3' similar to TR:Q9RIJ2 Q9RIJ2 BEX1 PROTEIN. ;, mRNA sequence.		

ACCESSION	BG228077	GR:12715592
VERSION	BG228077.1	
KEYWORDS	EST	
SOURCE	house mouse	
ORGANISM	Mus musculus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	
REFERENCE	1 (bases 1 to 587)	
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .	
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	
JOURNAL	Unpublished (1997)	
COMMENT	Contact: Robert Strausberg, Ph.D.	

Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:1393813

High quality sequence stop: 439.

FEATURES

source
Location/Qualifiers
1..587
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3513237"
/clone_lib="Soares_NMAX_maxillary_process"
/tissue_type="maxillary process"
/lab_host="DH10B (phage-resistant)"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site.1: NotI; Site.2: EcoRI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAGTGGAGCGCGCGCGCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 138 a 160 c 121 g 168 t

ORIGIN

alignment_scores:
Quality: 586.00 Length: 117
Ratio: 5.186 Gaps: 1
Percent Similarity: 96.581 Percent Identity: 95.726

alignment_block:

US-09-327-750D-30 x BG228077/rev ..

Align seg 1/1 to reverse of: BG228077 from: 1 to: 587

13 MetGluAsnAspHisGlnLysGluLysGluLysGluLysProGlnAs 29
587 ATGAGAATGACCATCAGAAAAAGGAGGAGGAAGAAATAGCCACAAGA 538

29 pThrIleArgArgGluPro.AlavalAlaLeuIleSerGluAlaGlyLys 45
537 TACCATPACAAGGAGGCCAGGCTGTGGCCCTGACCTCCGAGGCTGCCAAA 488

46 AspCysAlaProArgGlyGlyArgArgPheArgValArgGlnProI 62
487 AACTGTGCACCTTAGAGGAGCTCGCAGCGGTTCCGGGTTCCGAGCCCAT 438

62 eAlaHisTyrArgTrpAspLeuMetGlnArgValGlyGluProGlnGly 79
437 CGCTCACTATAGTGGACCTGATGCAGAGGGTTGGGAGCCCCAGGGAA 388

79 rgMetArgGluGluAsnValGlnArgPheGlyGlyAspValArgGlnLeu 95
387 GGATGAGAGGAGAACGTACAGAGGTTTGGGGGTGATGTGAGACAGCTC 338

96 MetGluLysLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSe 112
337 ATGGAGAACTGAGGAAAGGAGCTGAGCCACAGCCTCGGGCGGTTAG 288

112 rThrAspProProHisHisAspHisHisAspGluPheCysLeuMetPro 128
287 CACTGACCCGCTCATCATGACCACCATGATGAGTTTTCCTCATGCCCC 239

seq_name: gb_est1:BE291071

seq_documentation_block:

LOCUS BE291071 583 bp mRNA EST 13-JUL-2000
DEFINITION 601086311F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3500522 5',
mRNA sequence.
ACCESSION BE291071
VERSION BE291071.1 GI:9172545
KEYWORDS EST.

SOURCE

ORGANISM

house mouse.
Mus musculus
Mammalia; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 583)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC).
Unpublished (1999)

AUTHORS

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

JOURNAL

Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM8559 row: e column: 03

High quality sequence start: 5

High quality sequence stop: 503.

FEATURES

source

1..583
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3500522"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 165 a 136 c 201 g 81 t

ORIGIN

alignment_scores:

Quality: 542.50 Length: 120
Ratio: 4.801 Gaps: 2
Percent Similarity: 94.167 Percent Identity: 90.000

alignment_block:

US-09-327-750D-30 x BE291071 ..

Align seg 1/1 to: BE291071 from: 1 to: 583

1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspHi 17

208 ATCGAGTCCAAAGATCAAGCGGTGAAAAATCTCAACATGGAGATGACCA 257

17 sGlnLysLysGluLysGluLysProGlnAspThrIleArgArg 34

258 TCAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 307

34 luProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50

308 AGCCAGCTGTGGCCCTGACCTCGAGGCTGGCAAAACTGTGCACCTAGA 357

51 GlyGlyArgArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67

358 GGAGGTGCGCAGCGGTTCCGGGTTCGGCAGCCCATCTCCTATATAGT 407

67 pAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluAla 84

408 GGACCTGATCGACAGGGTTGGGAGCCCCCAGTGAAGGATGAGAGGAGA 457

84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100

458 ACGTACAGAGGTTAGGGGGTGATGTGAGACAGCTCATGGAGAGCTGAGG 507

```

101 GluArgGlnLeuSerHis...SerLeuArgAlaVal_SerThrAspProp 116
||||| ||||||| :::: ||||||| ::::::|||
508 GAAGGCTGTGAGCCACACACATCGAGCGGTTAGACAATGACCCGA 557
|||||
116 roHisHis 118
|||||
558 CTCATCAT 565

seq_name: gb_est2:BF608209

* seq_documentation_block:
LOCUS BF608209 650 bp mRNA EST 01-APR-2001
DEFINITION MY1_001088 Mouse 9-day fetus cDNA library ICRFp522 Mus musculus
cDNA clone ICRFp522C0878 5', mRNA sequence.
ACCESSION BF608209
VERSION BF608209.1 GI:13504638
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 650)
Yahyawi,M., Hennig,S., Neidhardt,L., Radelof,U., Hermann,B.G.,
Lehrach,H. and O'Brien,J.
Detection of a high number of novel genes in a 9-day mouse embryo
cDNA library normalised by oligonucleotide fingerprinting
JOURNAL Unpublished (2001)
COMMENT Contact: Hennig S
Laboratory 123, dept.Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Innestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennig@molgen.mpg.de
EST's are made from clones being representatives of clone clusters.
Clone clusters were calculated from oligonucleotide fingerprints.
PCR Primers
FORWARD: 5'-GAGCTATTCCAGACTAGTGA-3'
BACKWARD: 5'-TAATACGACTCATTATAGG-3'
Seq primer: 5'-ATTAGTGACACTATAG-3'
High quality sequence stop: 650.
Location/Qualifiers
1..650
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="ICRFp522C0878"
/clone_lib="Mouse 9-day fetus cDNA library ICRFp522"
/tissue_type="whole embryo"
/dev_stage="embryonic 9-day"
/lab_host="E.coli, XL1 blue"
/note="Vector: PSVSPORT1; Site_1: NotI; Site_2: SalI;
Library preparation by oligo_dT priming of RNA. Clones can
be ordered from the Resource Center in Berlin,
http://www.rzpd.de."
BASE COUNT 164 a 167 c 180 g 128 t 11 others
ORIGIN

alignment_scores:
Quality: 537.00 Length: 128
Ratio: 4.629 Gaps: 0
Percent similarity: 90.625 Percent identity: 79.688

alignment_block:
US-09-327-750d-30 x BF608209
Align seq 1/1 to: BF608209 from: 1 to: 650

1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspH1 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 ATGGATCCCAAGATCAAGCGGTAAGAAATCTCAACATGGACACACCA 255
|||||
17 sGlnLysLysGluGluLysProGlnAspThrIleArgArg 34

```

```

|:::|||||:|||||:|||||:|||||:|||||:|||||:|||||
256 TCACAAAAGGACCAAGCAAGCAAAAGCCCAAGATACCATCTTAATGG 305
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
34 luProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
306 AGCAGCTGTGGCCCTTACCTCCTAGCTAGCTAGCAAAAACCTGTGCACCTTGA 355
|||||
51 GlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
356 GNAAGTCGACAGGCGGTTCCGGGTCGGCAGCCCATCGCTTACTATANATC 405
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
67 pasLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluAla 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
406 GGACCTGATCCACAGGGTTGGGAGCCCATGGGAAGGATCAGACACACA 455
|||||
84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
456 ACGTACAGAGGCTCTGGGGGTGATGTGACAGAGCTCATGGAGAAACTCAGG 505
|||||
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPropProH1 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
506 GAAAGGCAGCTTACCCACAGCCTCGCGGCGGTAGCAGCTGACNCGCCTCA 555
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
556 TCATGACTACCATGATTAGTTTTCCTCATGCCCC 589
|||||
seq_name: gb_gss:AZ936393

```

```

seq_documentation_block:
LOCUS AZ936393 637 bp DNA GSS 26-APR-2001
DEFINITION ZM0193105F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M0193L05 F, DNA sequence.
ACCESSION AZ936393
VERSION AZ936393.1 GI:13794974
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 637)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenan,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.,
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0193 row: L column: 05
Seq primer: CGTTGTAAGACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 637.
Location/Qualifiers
1..637
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0193L05"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource

```


Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5'-GAGAGAGAGCGGCGCACTCGAGTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 5.0. Second strand cDNA
was prepared with the primer adapter of sequence[5',
CAGAGAGAGAGATCCAGAGTCAATTAAATTAATTAACCCCCCCCCC 3']. cDNA was
cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end:
XhoI. Host: SOLR.

FEATURES

Location/Qualifiers
1..480

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGD:MGI:1328321"
/db_xref="MGD:MGI:1900507"
/clone="1110004J10"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="18 days embryo"
c1..261

CDS

/note="putative"
/codon_start=1
/protein_id="BAB22786.1"
/db_xref="GI:12834097"
/translation="SGRNCAPRGRRRVRPROPIAHYRWDLMQRVGEPQGRMRRENVQ
RFGDVRQLMEKLRERQLSHSLRAVSTDPHPHHDHDFCLMP"
460..465
/note="putative"
480

polyA_signal

polyA_site

BASE COUNT 117 a 103 c 132 g 128 t
ORIGIN

alignment_scores:
Quality: 471.00 Length: 86
Ratio: 5.477 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.837

alignment_block:

US-09-327-750D-30 x AK003429 ..
Align seg 1/1 to: AK003429 from: 1 to: 480
43 AlaGlyLysAsnCysAlaProArgGlyArgArgPheArgVal 59
:::|||||
1 TCTGGCAAAACCTGTCACCTAGAGGAGTGCAGCGGTTCCGGGTCG 50
59 gGlnProIleAlaHisTyrArgTrpAspLeuMetGlnArgValGlyGluP 76
|||||
51 GCAGCCCATCGCTCACTATAGTGGACCTGATGCAGAGGGTTGGGGAGC 100
76 roGlnGlyArgMetArgGluGluAsnValClnArgPheGlyGlyAspVal 92
|||||
101 CCCAGGGAAGGATGAGAGAGGAGAACGTACAGAGGTTGGGGGTGATGTG 150
93 ArgGlnLeuMetGlnLysLeuArgGluArgGlnLeuSerHisSerLeuAr 109

|||||
151 AGACAGCTCATGGAGAAAGCTGAGGGAAAGGACAGCTGAGCCACAGCCTGCG 200
109 gAlaValSerThrAspProHisHisAspHisHisAspGluPheCysL 126
|||||
201 GCGGGTTAGCACTGACCCCGCCTCATCATGACCACCATGATGATTTTGCC 250
126 euMetPro 128
|||||
251 TCATGCCC 258

THIS PAGE BLANK (USPTO)


```

Align seg 1/1 to: US-08-381-881-1 from: 1 to: 1315
43 AlaGlyLysAsnCysAlaProArgGlyGlyArgArgPheArgValar 59
||||:||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
542 GCAAGCCAAACAGTCTCTCTGGCGCGCGCTCGCGCGCTCG 591
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
59 gGlnProIle.....AlaHisTyrArgTrp..... 67
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
592 CCGCCCATGCCGAGAGCGGCACCATTTGTCATCCGCGCAGTACCGCAC 641
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
68 ..AspLeuMetGlnArgValGlyGluPro.....GlnGlyArgMet 80
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
642 CAGACCTGCAGCATCGAGAGCCGCGCGCTTACATCAAGGTGCGTGA 691
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
81 ArgGluGluAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetG1 97
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
692 CCTGCCAC.....GATCTCCAAGAGCGCGCTGA 720
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
97 uLysLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThra 114
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
721 AGAAGCGCGGCGAGCTGGCGGGCGGCGCTCGCTTCGATATCAAGCTGAAG 770
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
114 spProHisHisAspHisAspGluPheCysLeu 126
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
771 GACTGCCCGACCGTCAACACTCTCAAGCTGTACTT 808
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-281-221-1

seq_documentation_block:
; Sequence 1, Application US/09281221
; Patent No. 6284256
; GENERAL INFORMATION:
; APPLICANT: SAKELAKUL, PAUL H.,M.
; TITLE OF INVENTION: BORDETTELLA BRONCHISEPTICA VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVENS, DAVIS, MILLER & MOSHER
; STREET: 515 NORTH WASHINGTON STREET.
; CITY: ALEXANDRIA
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/281,221
; FILING DATE: 30-Mar-1999
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/381,881
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: POULOS III, JAMES A.
; REGISTRATION NUMBER: 31,714
; REFERENCE/DOCKET NUMBER: TPP 29685
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703 549-7200
; TELEFAX: 703 528-5313
; TELEX: 44-0704 "STEVENS"
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1315 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO

```

```

; ORIGINAL SOURCE:
; ORGANISM: Bordetella bronchiseptica
; STRAIN: 401
; IMMEDIATE SOURCE:
; CLONE: E coli PC2495(pIVB3-420)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..539
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 540..1142
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1143..1315
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-281-221-1

alignment_scores:
Quality: 93.50 Length: 96
Ratio: 1.908 Gaps: 4
Percent Similarity: 51.042 Percent Identity: 33.333

alignment_block:
US-09-327-750D-30 x US-09-281-221-1 ..
Align seg 1/1 to: US-09-281-221-1 from: 1 to: 1315
43 AlaGlyLysAsnCysAlaProArgGlyGlyArgArgPheArgValar 59
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
542 GCAAGCCAAACAGTCTCTCTGGCGCGCGCTCGCGCGCTCG 591
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
59 gGlnProIle.....AlaHisTyrArgTrp..... 67
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
592 CCGCCCATGCCGAGAGCGGCACCATTTGTCATCCGCGCAGTACCGCAC 641
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
68 ..AspLeuMetGlnArgValGlyGluPro.....GlnGlyArgMet 80
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
642 CAGACCTGCAGCATCGAGAGCCGCGCGCTTACATCAAGGTGCGTGA 691
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
81 ArgGluGluAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetG1 97
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
692 CCTGCCAC.....GATCTCCAAGAGCGCGCTGA 720
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
97 uLysLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThra 114
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
721 AGAAGCGCGGCGAGCTGGCGGGCGGCGCTCGCTTCGATATCAAGCTGAAG 770
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
114 spProHisHisAspHisAspGluPheCysLeu 126
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||
771 GACTGCCCGACCGTCAACACTCTCAAGCTGTACTT 808
||||:||||| ||| ||| ||| ||| ||| ||| ||| |||

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-931-999-4

seq_documentation_block:
; Sequence 4, Application US/08931999
; Patent No. 6043219
; GENERAL INFORMATION:
; APPLICANT: Tandolo, John J.
; APPLICANT: Crupper, Scott S.
; TITLE OF INVENTION: Broad Spectrum Chemotherapeutic Peptide
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hovey, Williams, Timmons & Collins
; STREET: 2405 Grand Boulevard, Suite 400
; CITY: Kansas City
; STATE: Missouri
; COUNTRY: U.S.A.
; ZIP: 64108
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; GENERAL INFORMATION:
;
; APPLICANT:
; TITLE OF INVENTION: Process for producing anthracyclines
; TITLE OF INVENTION: and intermediates thereof
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKown, Edwards and Lenahan
; STREET: 1200 G Street, Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,740A
; FILING DATE: 27-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FI95/00537
; FILING DATE: 30-SEP-1995
; APPLICATION NUMBER: FI 944556
; FILING DATE: 30-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: H. Thomas Anderton, Jr.
; REGISTRATION NUMBER: 40,895
; REFERENCE/DOCKET NUMBER: 1574/43419
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; STRAIN: Streptomyces nogalater ATCC 27451
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 359..1651
; OTHER INFORMATION: /note= "ORF1"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2937..3197
; OTHER INFORMATION: /note= "ORF3"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1648..1651
; OTHER INFORMATION: /note= "overlapping sequence in
; OTHER INFORMATION: ORF1 and ORF2"
;
US-08-809-740A-1

alignment_scores.
Quality: 85.00 Length: 127
Ratio: 1.197 Gaps: 6
Percent Similarity: 55.906 Percent Identity: 29.921

alignment_block:
US-09-327-750D-30 x US-08-809-740A-1/rev ..

Align seg l/1 to reverse of: US-08-809-740A-1 from: 1 to: 3252

16 AspHisGluLysLysGluGluLysGluGluLysProGlnAspThr...IL 31
:::||||| :::: ::: :||| |||:::
2772 GAGCACCAGTCCAGTCGTACTCCGGCGCGGGAGAGATTGACGG 2723
:::||||| :::: ::: :||| |||:::

31 eargArgGluProAlaValAlaLeuIlleSerGlu..... 42
|||||::: |::|::|::|
2722 TCGGGGGGACGAGCGCTCCGCAGCGCCGCCAGCGCGTCCGCCACGTGG 2673
:::||||| :::: ::: :||| |||:::

```

COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,599B
; FILING DATE: 21-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Stephen A.
; REGISTRATION NUMBER: 38,609
; REFERENCE/DOCKET NUMBER: 1604-123A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-624-1589
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Optimized cDNA"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..621
; US-08-426-599B-3

```

```

alignment_scores:
  Quality: 83.50      Length: 96
  Ratio: 1.670       Gaps: 6
  Percent Similarity: 52.083      Percent Identity: 34.375

```

```

alignment_block:
US-09-327-750D-30 x US-08-426-599B-3

```

Align seg 1/1 to: US-08-426-599B-3 from: 1 to: 624

```

33 ArgGluProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaPr 49
||||:|||||
15 CGTCAACG.....CTGCTCTACACCGTCTGGCGCGCGCTCCNA.. 56
49 oArgGlyArgArgArgPheArgValArgGlnProIleAlaHisTyrA 66
||||:|||||
57 .....GGCGAGCGGCGCTCGACGCGCGCGCGCGCCAGCTCAACTCC 99
66 rGTrpAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGlu 82
|||||
100 GG.....CGAGTCTGGACCATCAACGTCGAGCCCGGCAC 134
83 GluAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLe 99
|||||
135 CAGAGCGCGGCAAGATCTGGGCGCGGCAC...CGACTCTACTTCGACGACT 181
99 uArgGluArgGlnLeuSerHis..... 106
182 CCGCGCGCGCATCTGCGGACCGGCAGCTGCGGCGCGCTCTCCAGTGC 231
107 ..SerLeuArgAlaValSerThrAspProHisHis 118
|||||
232 AAGGCTTCGG.....CCGCGCGCGCCACCA 257

```

seq_name: /cqn2_6/ptodata/2/ina/6A_COMB.seq:US-08-506-553C-6

```

seq_documentation_block:
; Sequence 6, Application US/08506553C
; Patent No. 6120989
; GENERAL INFORMATION:
; APPLICANT: Vornhagen, Rolf; Hinderer, Walter; Sonneborn, Han-H.;
; APPLICANT: Plachtner, Bodo; and Jahn, Gerhard
; TITLE OF INVENTION: ISOLATED HUMAN CYTOMEGALOVIRUS
; TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF

```

```

; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/506,553C
; FILING DATE: 07/25/95
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LEDER 202-PFF/MAS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 607 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-506-553C-6

```

```

alignment_scores:
  Quality: 82.50      Length: 117
  Ratio: 1.650       Gaps: 4
  Percent Similarity: 42.735      Percent Identity: 26.496

```

alignment_block:

US-09-327-750D-30 x US-08-506-553C-6

Align seg 1/1 to: US-08-506-553C-6 from: 1 to: 607

```

50 ArgGlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrA 66
||||:|||||
107 CGGTCTGGAGAGGTGGCGCGGTGGCGCAAGAA...GCACGACGCG 153
66 gTrpAspLeuMetGlnArgValGlyGluProGln..... 77
|||||
154 GTGGCGCGGTGTTCGGTACGCGAAATAGTAGCGGTGGCGCGGC 203
77 ..... 77
204 GTGTATCATCACCGTCTTCTCTCAAGAAAAATACGACGACGACAA 253
78 .....GlyArgMetArgGlu 83
|||||
254 GATCACCAGCTACCTGACGTCCAAAGGTGGATCGGCGCGCGGAGGAG 303
83 uAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuA 100
||||:|||||
304 GAGGAGCGCGCGTTTGGATCGCAACTCCGCAATTACTTCAACGACGCG 353
100 rg.....GluArgGlnLeuSer..... 105
|||||
354 AAGAGAGAGACGACACCGAGGATTCGTACGTTCCAGTTCCGCCATAA 403
106 HisSerLeuArgAlaValSerThrAspProHisHisAspHisHis 122
|||||

```

404 CACCAAGAAGCAAAAGTCGGCAAGATCCTGAAGACGACGACGGCATGA 453

122 p 122

454 A 454

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-030-995-1

seq_documentation_block:

; Sequence 1, Application US/09030995

; Patent No. 6235519

; GENERAL INFORMATION:

; APPLICANT: Wang, Yongzhao

; APPLICANT: Childs, John D.

; APPLICANT: Squires, Charles H.

; TITLE OF INVENTION: A gene involved in thiophene

; FILE OF INVENTION: Biotransformation from No. 6235519ardia Asteroides KGB1

; FILE REFERENCE: EBC97-05

; CURRENT APPLICATION NUMBER: US/09/030,995

; CURRENT FILING DATE: 1998-02-26

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1

; LENGTH: 1365

; TYPE: DNA

; ORGANISM: No. 6235519ardia asteroides

US-09-030-995-1

alignment_scores:

Quality: 82.50 Length: 123

Ratio: 1.231 Gaps: 8

Percent Similarity: 54.472 Percent Identity: 30.894

alignment_block:

US-09-327-750D-30 x US-09-030-995-1 ..

Align seg 1/1 to: US-09-030-995-1 from: 1 to: 1365

27 ProGlnAspThrIleArgArg...GluProAlaValAlaLeuIleSerG1 42

908 CCTATGTCACAAATCCGCGCGTGTCTACATCGGCATCGACGACGACGTGG 957

42 uAlaGlyLysAsnCysAla...ProArgGlyGlyArgArg..... 55

958 TTCACCTTCACATGTCGACGCCCGCGCTGTGGCGCGCGACGTGAT 1007

56PheArgValArgGlnProIleAlaHisTyArgTrpAspLeuMet 70

1008 CCTGGCGCGCATCGCCCTGCCGCGGCAAGCAGGAGATGTCGCCGATG 1057

71 GlnArgValGlyGlu.....Pro.GlnGlyArgMetArgGluG 83

1058 TCGAGGCGCGCGTCCGCCCGCGGAGACCGCGCAAGGACGACTACGACGCG 1107

83 luAsnValGln.....ArgPheGlyGlyAspValArgGlnLeuMet 96

1108 ATCCGCTATCAGGGCGGACTAGTCAAGAGCTGATCCCGGAGACCGACTA 1157

97 GluLysLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerTh 113

1158 TCCGAGCTTCGACGTCGACGCGGCCAACGAAGCCTCTTCGAGTGAAGA 1207

113 rAsp.....ProPro..... 116

1208 AGCACAGAAGAAGAACATCATGAGTTCGCGCCACACACTCCTATCGCTCG 1257

117 ..HisHisAspHisHis 121

1258 GTCATCACCGGCACCAT 1274

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-403-852D-1

seq_documentation_block:

; Sequence 1, Application US/08403852D

; Patent No. 5891655

; GENERAL INFORMATION:

; APPLICANT: Blanc, Veronique

; APPLICANT: Blanc, Francis

; APPLICANT: Crouzet, Joel

; APPLICANT: Jacques, Nathalie

; APPLICANT: Lacroix, Patricia

; APPLICANT: Thibaut, Denis

; APPLICANT: Zagorec, Monique

; APPLICANT: Debussche, Laurent

; TITLE OF INVENTION: Polypeptides Involved In The

; TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fionegan, Henderson, Farabow, Garrett & Dunner

; STREET: 1300 I Street, N.W., Suite 700

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20005-3315

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/403,852D

; FILING DATE: 10-MAY-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/FR 93/00923

; FILING DATE: 25-SEP-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: FR 92/11441

; FILING DATE: 25-SEP-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Meyers, Kenneth J.

; REGISTRATION NUMBER: 25,146

; REFERENCE/DOCKET NUMBER: 03806.0054-00000

; TELEPHONE: (202) 408-4000

; TELEFAX: (202) 408-4400

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5392 base pairs

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: S.pristinaespiralis

US-08-403-852D-1

alignment_scores:

Quality: 81.50 Length: 118

Ratio: 1.430 Gaps: 5

Percent Similarity: 48.305 Percent Identity: 27.966

alignment_block:

US-09-327-750D-30 x US-08-403-852D-1 ..

Align seg 1/1 to: US-08-403-852D-1 from: 1 to: 5392

17 HisGlnLysLysGluGluLysGluLysProGlnAspThrIleArgAr 33

4800 CATCGTGTCACCGGTTCATCGCCACGACCATCTGATGGTCTTCCCCG 4849


```
33 g.....GluProAla.....ValAlaL 39
|
4850 GCGGTTTCGGGATGACGTATCCCGACAGCTCGCTCATGTCTCGCTC 4899
|
39 eutleSerGluAlaGlyLysAsnCysAlaPro.....ArgGlyGlyArg 53
|
4900 TCCTTCTGTCAGCAGCTCGAGGTGCGCGCGGGGAGTGGCGGACAA 4949
|
54 ArgArgPheArgValArgGlnProIleAlaHisTyrArgTrpAspLeuMe 70
|
4950 CGTCGCTTCGGCTCGCGCGCTCGCGCTCAC..... 4982
|
70 tGlnArgValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnA 87
|
4983 .....CCCCAGCTGTCGCGCGCTCGCGCTCAC..... 5019
|
87 rgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArgGln 103
|
5020 GA.....GTACGAGGTCTGGCTCAAGGAACACGCGCTCGAC 5057
|
104 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 120
|
5058 CGGCGCGCTCTGCTCTCCACCGAGCGGACAGCGCGCTTCATGTGCA 5107
|
120 shis 121
|
5108 TCAC 5111
```

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-510-646B-1

```
seq_documentation_block:
; Sequence 1, Application US/08510646B
; Patent No. 6077699
; GENERAL INFORMATION:
; APPLICANT: Blanc, Veronique
; APPLICANT: Blanche, Francis
; APPLICANT: Crouzet, Joel
; APPLICANT: Jacques, Nathalie
; APPLICANT: Lacroix, Patricia
; APPLICANT: Thibaut, Denis
; APPLICANT: Zagorec, Monique
; APPLICANT: Debussche, Laurent
; APPLICANT: De Crecy-Lagard, Valerie
; TITLE OF INVENTION: Polypeptides Involved In The
; TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
; TITLE OF INVENTION: Coding For These Polypeptides And Their Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/510,646B
; FILING DATE: 03-AUG-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,852
; FILING DATE: 10-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR 93/00923
; FILING DATE: 25-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/11441
; FILING DATE: 25-SEP-1992
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03806.0054-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5392 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: S.pristinaespiralis
;
US-08-510-646B-1
```

```
alignment_scores:
Quality: 81.50 Length: 118
Ratio: 1.430 Gaps: 5
Percent Similarity: 48.305 Percent Identity: 27.966
alignment_block:
US-09-327-750D-30 x US-08-510-646B-1
```

```
Align seg 1/1 to: US-08-510-646B-1 from: 1 to: 5392
17 HisGlnTylsGluGluLysGluGluLysProGlnAspThrIleArgAr 33
|
4800 CATCGCTGTCAACCGGTTCATCGCCACCGACCATCTCATGTCTCCCG 4849
|
33 g.....GluProAla.....ValAlaL 39
|
4850 GCGGTTTCGGGATGACGTATCCCGACAGCTCGCTCATGTCTCGCTC 4899
|
39 eutleSerGluAlaGlyLysAsnCysAlaPro.....ArgGlyGlyArg 53
|
4900 TCCTTCTGTCAGCAGCTCGAGGTGCGCGCGGGGAGTGGCGGACAA 4949
|
54 ArgArgPheArgValArgGlnProIleAlaHisTyrArgTrpAspLeuMe 70
|
4950 CGTCGCTTCGGCTCGCGCGCTCGCGCTCAC..... 4982
|
70 tGlnArgValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnA 87
|
4983 .....CCCCAGCTGTCGCGCGCTCGCGCTCAC..... 5019
|
87 rgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArgGln 103
|
5020 GA.....GTACGAGGTCTGGCTCAAGGAACACGCGCTCGAC 5057
|
104 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 120
|
5058 CGGCGCGCTCTGCTCTCCACCGAGCGGACAGCGCGCTTCATGTGCA 5107
|
120 shis 121
|
5108 TCAC 5111
```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-231-818-1

```
seq_documentation_block:
; Sequence 1, Application US/09231818
; Patent No. 6171846
; GENERAL INFORMATION:
; APPLICANT: Blanc, Veronique
; APPLICANT: Blanche, Francis
; APPLICANT: Crouzet, Joel
; APPLICANT: Jacques, Nathalie
; APPLICANT: Lacroix, Patricia
```



```

981 AGAAAAAGGTATACATATGATAAATGGGAAAGTTACACAGAAAGATG 1030
      ::::: ::::: ::::: ::::: ::::: :::::
78  lYArgMetArgGluGluAsnValGlnArgPheGlyAspValArgGln 94
      ::::: ::::: ::::: ::::: ::::: :::::
1031 AAGAAITGGAGGACAGTGTGCCAGCAT.....GGGAGAGTACATGAG 1074
      ::::: ::::: ::::: ::::: ::::: :::::
95  LeuMetGluLysLeuArgGluArgGlnLeuSerHisSerLeuArgAlaVa 111
      ::::: ::::: ::::: ::::: ::::: :::::
1075 ACGATG...AAGCAAGGCTAAGCGAGCTGGATAAGCAGCAGCCAC 1121
      ::::: :::::
111 lSer 112
      :::::
1122 AGCC 1125

```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-393-985-17

```

seq_documentation_block:
; Sequence 17, Application US/08393985
; Patent No. 5693476
; GENERAL INFORMATION:
; APPLICANT: Scheller, Richard H.
; TITLE OF INVENTION: Methods and Compositions for Modulation
;   OF INVENTION: of Vesicular Release
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,985
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 8600-0152
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2040 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Mouse SNAP-25 (GenBank M22012)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 164..784
US-08-393-985-17

```

```

alignment_scores:
  Quality: 81.00      Length: 125
  Ratio: 1.110      Gaps: 6
Percent Similarity: 58.400      Percent Identity: 25.600

```

alignment_block:

US-09-327-750d-30 x US-08-393-985-17

```

Align seg 1/1 to: US-08-393-985-17 from: 1 to: 2040

2  GluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspHisG1 18
      ||||| ||||| ||||| ::::: ::::: |||||
275 GAGAGTAAAGATCTGTCATPCAGGACTTTGGTTATGTTGATGAGCAAG 324
      ||||| ||||| ||||| ::::: ::::: |||||
18  nLysLysGluGluGluGluLysProGlnAspThrIleArgArgGluP 35
      ||||| ||||| ||||| ::::: ::::: |||||
325 CGAACAACCTGGAACGATTGAGGAGGATGGACCAATCAATGAAGAT. 373
      ||||| ||||| ||||| ::::: ::::: |||||
35  roAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArgGly 51
      ::::: ::::: ||||| ||||| ::::: |||||
374 .....ATGAAAGAACGACAAAGAATTGTGACGACCTAGGA 409
      ::::: :::::
52  GlyArgArgArgPhe.....ArgValar 59
      ::::: :::::
410 .....AAATTCGCGGGCTTTGTGTGTGTCCTGTAAACAGCTTAA 450
      ::::: :::::
59  gGlnProIleAlaHisTyrArg...TirAspLeuMetGlnArg.....V 73
      ::::: ||||| ::::: |||||
451 ATCCAGTGATGCTTACAAAAAGCCTGGGCAATAATCAGGATGGAGTAG 500
      ::::: ||||| ::::: |||||
73  aGlyGluProGlnGlyArgMetArgGluGluAsnValGlnArg..... 87
      ||||| ::::: ::::: |||||
501 TGGCCAGCCAGCCTGCCCTGTGTGTGATGACGGGAGCAGATGGCCATC 550
      ||||| ::::: ::::: |||||
88  PheGlyGlyAspValArgGlnLeuMetGluLysLeuArgGluArgGlnLe 104
      ||||| ::::: ::::: |||||
551 AGTGGTGGCTTCATCCGAGGGTAACAACGATGCCCGGAAATGAAT 600
      ||||| ::::: ::::: |||||
104 userHisserLeuArgAlaValSer 112
      ::::: ::::: |||||
601 GGATGAAACCTAGACAGGTGAGC 625

```

seq_name: /cgn2_5/ptodata/2/ina/5A_COMB.seq:US-07-951-715A-6

```

seq_documentation_block:
; Sequence 6, Application US/07951715A
; Patent No. 5625136
; GENERAL INFORMATION:
; APPLICANT: Koziele, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merilin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
;   OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/951,715A

```

```
seq_documentation_block:
: Sequence 6, Application US/08459448A
: Patent No. 5859336
: GENERAL INFORMATION:
: APPLICANT: Koziel, Michael G.
: APPLICANT: Desai, Nalini M.
: APPLICANT: Lewis, Kelly S.
: APPLICANT: Kramer, Vance C.
: APPLICANT: Warren, Gregory W.
: APPLICANT: Eviola, Stephen V.
: APPLICANT: Crossland, Lyle D.
: APPLICANT: Wright, Martha S.
: APPLICANT: Merlin, Ellis J.
: APPLICANT: Launis, Karen L.
: APPLICANT: Rothstein, Steven J.
: APPLICANT: Bowman, Cindy G.
: APPLICANT: Dawson, John L.
: APPLICANT: Dunder, Erik M.
: APPLICANT: Pace, Gary M.
: APPLICANT: Suttie, Janet L.
: TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
: INSECTICIDAL ACTIVITY IN MAIZE
: NUMBER OF SEQUENCES: 94
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 5859336artis Corporation
: STREET: Patent & Trademark Dept., 520 White Plains
: STREET: Rd., POB 2005
: CITY: Tarrytown
: STATE: New York
: COUNTRY: USA
: ZIP: 10591-9005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/459,448A
: FILING DATE: 02-JUN-1995
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/951,715
: FILING DATE: 25-SEP-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/772,027
: FILING DATE: 04-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Pace, Gary M.
: REGISTRATION NUMBER: 40403
: REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV4
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919)541-8582
: TELEFAX: (919)541-8689
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3624 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: other nucleic acid
: DESCRIPTION: /desc = "Synthetic DNA"
: HYPOTHETICAL: NO
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..3621
: OTHER INFORMATION: /product= "Full-length, maize
: OTHER INFORMATION: optimized cryiB"
: OTHER INFORMATION: /note= "Disclosed in Figure 6."
US-07-951-715A-6

alignment_scores:
: Quality: 81.00 Length: 121
: Ratio: 1.397 Gaps: 4
: Percent Similarity: 47.934 Percent Identity: 23.967

alignment_block:
US-09-327-750D-30 x US-07-951-715A-6 ..
Align seg 1/1 to: US-07-951-715A-6 from: 1 to: 3624

14 GluAsnAspHisGlnIysLysGluGluLysGluLysProGlnAspTh 30
||||:||||| :||||:||||| :||||:|||||
1026 GAACACGAGCAGCCAGCGCGCCACCAACACGACGATCAACCGGTGACCC 1075
30 rIleArgArgGluProAlaValAlaLeuIleSerGluAlaGlyLysAsnC 47
||||:||||| :||||:||||| :||||:|||||
1076 TGGCTTCGCGCGCGGCGGTGTACCGCACCGAGAGCTACGCGGCGTG 1125
47 ysAla.....ProArgGlyGlyArgArgArg 55
:||||| :||||| :|||||
1126 CTGCTGGGGGCATCTACCTGGAGCCCATCCACGGCGTGCCACCGTGG 1175
56 PheArgValArgGlnProIleAlaHisTyrArgTrpAspLeuMetGlnAr 72
:||||:||||| :||||:||||| :||||:|||||
1176 CTTCAACTTCACCAACCCCGGACGACATCAGCGA.....CCGCG 1213
72 gValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnArgPheG 89
| :||| :||| :||| :||| :|||
1214 GCACCGCCAACTACAGCAGCCCTACGAGAGCC..... 1247
89 lyGlyAspValArgGlnLeuMetGluLysLeuArgGluArg..... 102
:||||| :||||| :|||||
1248 .....CGCGCTGCAGCTGAAGGACGAGCAGCGAGCTGCCGCC 1286
103 .....GlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
: : : : :
1287 CGAGACGACGCGCGCCCACTACGAGAGCTACAGCCACCGCTGAGCC 1336
117 sHisAspHisHis 121
||| |||||
1337 ACATCGGCATCAT 1349

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-459-448A-6
alignment_scores:
```


OM of: US-09-327-750D-30 to: N_Geneseq_1101:* out_format : pfs
Date: Mar 11, 2002 3:42 PM
About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=cgn2_1/USPTO.spool/US09327750/runat_11032002_101154_20362/app_query.fasta_1.1472
-DB=N_Geneseq_1101 -Qfmt=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blom62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USR=US09327750.cGNI_1_330 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-327-750D-30
Query length: 128
Database: N_Geneseq_1101:*
Database sequences: 930621
Database length: 428662619
Search time (sec): 355.560000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143 +		450.50	857.22	792	1.1e-39
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH75810 +		450.00	856.56	767	1.2e-39
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH158581 +		450.00	855.45	862	1.4e-39
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF59611 +		450.00	855.06	898	1.5e-39
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH03880 +		437.00	832.87	662	2.6e-38
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH03879 +		437.00	832.36	698	2.8e-38
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH160367 +		437.00	830.40	858	3.5e-38
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH10889 +		429.50	817.97	692	1.7e-37
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517 +		171.50	317.52	891	1.3e-09
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517 +		170.00	314.90	865	1.8e-09
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13750 +		170.00	311.56	1229	2.8e-09
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH85548 +		170.00	310.56	1364	3.2e-09
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF23528 +		160.00	297.62	700	1.7e-08
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAF21748 +		143.50	263.19	917	1.4e-06
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01005 +		119.50	222.05	532	0.0083
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH87147 +		100.50	195.34	187	0.0084
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH06183 +		99.00	184.71	421	0.0327
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH06183 +		99.00	183.93	457	0.0361
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01004 +		93.50	172.74	485	0.1317
/SID2/gcgdata/geneseq/geneseq/NA1989.DAT:AAH090292 +		93.50	167.46	844	0.2985
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH06184 +		91.50	169.58	451	0.2277
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH16713 +		91.50	155.07	2067	1.46
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH82586 +		88.50	164.33	426	0.4464
/SID2/gcgdata/geneseq/geneseq/NA1997.DAT:AAH74547 +		87.50	135.09	7481	18.98
/SID2/gcgdata/geneseq/geneseq/NA1998.DAT:AAH21511 +		86.50	134.13	6755	21.46
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH14635 +		86.00	140.28	3201	9.75
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH75924 +		86.00	136.57	4724	15.69
/SID2/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ00461 +		85.50	146.85	1452	4.20
/SID2/gcgdata/geneseq/geneseq/NA1998.DAT:AAH61594 +		85.00	151.70	789	2.26
/SID2/gcgdata/geneseq/geneseq/NA1996.DAT:AAH25480 +		85.00	138.20	3252	12.73
/SID2/gcgdata/geneseq/geneseq/NA1996.DAT:AAH06312 +		83.50	151.04	624	2.45
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH87834 +		83.00	151.07	562	2.44
/SID2/gcgdata/geneseq/geneseq/NA1999.DAT:AAH11192 +		83.00	136.91	2484	15.03
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH51286 +		83.00	135.94	2751	17.03
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH65399 +		82.50	149.48	600	3.00
/SID2/gcgdata/geneseq/geneseq/NA1995.DAT:AAH03508 +		82.50	149.37	607	3.04
/SID2/gcgdata/geneseq/geneseq/NA1999.DAT:AAH21172 +		82.50	141.65	1365	8.18
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH65400 +		82.50	136.82	2265	15.20
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH68525 +		82.50	88.79	349980	7.2e+03
/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH47167 +		82.00	131.21	3688	31.20

/SID2/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ53884 + 81.50 143.82 6.20
/SID2/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ64201 + 81.50 126.63 56.18
/SID2/gcgdata/geneseq/geneseq/NA1999.DAT:AAH40054 + 81.00 143.24 6.67
/SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH00370 + 81.00 142.48 7.35
/SID2/gcgdata/geneseq/geneseq/NA1998.DAT:AAV57358 + 81.00 134.92 19.38

seq_name: /SID2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143

seq_documentation_block:

ID AAH45143 standard; cDNA; 792 BP.
XX
AC AAH45143;
XX
DT 07-SEP-2001 (first entry)
XX
DE Human brain expressed X-linked protein, hBex, coding sequence.
XX
KW Human; brain expressed X-linked protein; cytostatic; auditory; nontropic;
KW hBex; dysembryoplasia; hereditary disease; cancer; tumour; deafness;
KW X-chromosome-binding mental retardation; lissencephalous disease; ss.
XX
OS Homo sapiens.
XX
PN WO200140286-A1.
XX
PD 07-JUN-2001.
XX
PF 27-NOV-2000; 2000WO-CN00502.
XX
PR 30-NOV-1999; 99CN-0124179.
XX
PA (BIOR-) BIORAD GENE DEV LTD SHANGHAI.

PI Mao Y, Xie V;
XX
DR WPI; 2001-397944/42.
XX
P-PSDB; AAB99224.
XX
PT Isolated human brain-expressed X-linked polypeptide used to diagnose
PT and treat of dysembryoplasia, hereditary diseases, cancer, tumor,
PT deafness and X-chromosome-binding mental retardation -
XX
PS Claim 5; Page 22; 30pp; Chinese.
XX

CC The present sequence is the coding sequence for a human brain-expressed
CC X-linked protein (hBex). hBex and its coding sequence are useful in the
CC diagnosis and treatment of dysembryoplasia, hereditary diseases, cancer,
CC tumours, deafness, X-chromosome-binding mental retardation and
CC lissencephalous disease. hBex is also useful for screening mimics,
CC agonists, or inhibitors, and in peptide fingerprinting identification.
CC hBex coding sequence can be used as primers or probes, or in producing
CC gene chips or microarrays.
XX

SQ Sequence 792 BP; 214 A; 172 C; 219 G; 187 T; 0 other;

alignment_scores:

Quality: 450.50 Length: 129
Ratio: 3.952 Gaps: 3
Percent Similarity: 88.372 Percent Identity: 68.992

alignment_block:

US-09-327-750D-30 x AAH45143 ..

Align seg 1/1 to: AAH45143 from: 1 to: 792

1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16

173 ATGGAGTCCAAAGAAACTAGCAGTGAACAGTCTCAGCATGCAAAATGC 222

16 pHisGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArga 33

223 CAACCAAGAAATGAAGAAAGGAG.....CAAGTTGCTAATAAAG 263


```
Query Match          4.9%; Score 34.4; DB 4; Length 1247;
Best Local Similarity 59.0%; Pred. No. 0.72;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 229 gacaggaacacccgcctgtggaggagggtgagggccaccagctgtgcaaacacaaca 288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1100 GACTGTACCATTCGACTCTGGCTGGGTGACAGACGACGACCCCTGTCTCAAAAAAGAAAA 1041

Qy 289 acaacaacacacaccataaccacaacaccaccaggagg 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1040 ACAAACAAAAACAAAAAACGCTACCTGCCTGAGCAGG 1001

RESULT 14
US-08-787-739-58/c
; Sequence 58, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2501 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: DNA (genomic)
; DESCRIPTION: MN genomic region between pMN1 and
; DESCRIPTION: Bd3 assumed to contain regulatory
; DESCRIPTION: elements
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-787-739-58

Query Match          4.9%; Score 34.4; DB 3; Length 2501;
Best Local Similarity 59.0%; Pred. No. 1.1;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 229 gacaggaacacccgcctgtggaggagggtgagggccaccagctgtgcaaacacaaca 288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2421 GACTGTACCATTCGACTCTGGCTGGGTGACAGACGACGACCCCTGTCTCAAAAAAGAAAA 2362

Qy 289 acaacaacacacaccataaccacaacaccaccaggagg 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2361 ACAAACAAAAACAAAAAACGCTACCTGCCTGAGCAGG 2322

RESULT 15
US-09-178-115-58/c
; Sequence 58, Application US/09178115
; Patent No. 6297041
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.5A
; CURRENT APPLICATION NUMBER: US/09/178,115
; EARLIER APPLICATION NUMBER: 09/177,776
; EARLIER FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 08/787,739
; EARLIER FILING DATE: 1997-01-24
; EARLIER APPLICATION NUMBER: 08/485,049
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/486,756
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/477,504
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/481,658
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,862
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,863
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/487,077
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/260,190
; EARLIER FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PV-709-92
; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 2501
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2501)
US-09-178-115-58

Query Match          4.9%; Score 34.4; DB 4; Length 2501;
```



```

; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 110
; LENGTH: 1247
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1247)
US-09-178-115-110

Query Match      4.9%; Score 34.4; DB 4; Length 1247;
Best Local Similarity 59.0%; Pred. No. 0.72;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps

QY 229 gacggaacaccgccctgtggaggagggtggggccaccagcctgtcgaacaacaaca 288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1100 GACTGTACCATTGCATCTCGGCTGGTGACAGAGGACGACCCCTGTCTCAAAAAAGAAA 1041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 289 acaacacaccacaccataaccacacacccaccgcgaagagg 328.
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1040 AAAAAACAAAAACAAAAACGCTACCTGCCCTGAGCAGG 1001
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-09-177-776-110/c
; Sequence 110, Application US/09177776A
; Patent No. 6297051
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.5A
; CURRENT APPLICATION NUMBER: US/09/177,776A
; EARLIER FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 08/787,739
; EARLIER FILING DATE: 1997-01-24
; EARLIER APPLICATION NUMBER: 08/485,049
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/486,756
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/477,504
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/481,658
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,862
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,863
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/487,077
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/260,190
; EARLIER FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PV-709-92
; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 110
; LENGTH: 1247
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1247)
US-09-177-776-110

```

TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
TITLE OF INVENTION: SPECIES INFECTIONS
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETERS, VERNY, JONES & BIKSA
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-1840
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,361B
FILING DATE: 12-SEP-1997
CLASSIFICATION:

APPLICATION NUMBER: US 80/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Verny, Hana

REGISTRATION NUMBER: 30,318
REFERENCE/DOCKET NUMBER: 480.76-1 (HV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-1677

```

/ INFORMATION FOR SEQUENCE ANALYSIS:
/
/ SEQUENCE CHARACTERISTICS:
/   LENGTH: 5318 base pairs
/   TYPE: nucleic acid
/   STRANDEDNESS: double
/   TOPOLOGY: linear
/   MOLECULE TYPE: DNA (genomic)
/ US-08-928-361B-3

```

```

Query Match      5.0%; Score 34.8; DB 3; Length 5310;
Best Local Similarity 84.8%; Pred. NO. 1.3;
Matches 39; Conservative 0; Mismatches 7; Indels 0

QY 279 aacaacaacaacaacacacacataaccacacaccaccacgaa 324
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 873 AACTACACCAACAACCAACCAACCAACCAACTACCAAGAA 918

RESULT 11

```

US-08-692-922-1/c ; Sequence 1, Application US/08692922
; Patent No. 6277592
; GENERAL INFORMATION:
; APPLICANT: Bidwell, Christopher

TITLE OF INVENTION: PORCINE LEPTIN PROTEIN, NUCLEIC ACID
 TITLE OF INVENTION: SEQUENCES CODING THEREFOR AND USES THEREOF
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:

ADDRESSEE: WHYTE HIRSCHBOECK DUDER S.C.
STREET: Suite 2100 111 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin

```

ZIP: 33202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(289)
OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

Query Match
Best Local Similarity 5.3%; Score 36.8; DB 4; Length 289;
Matches 15; Conservative 99; Mismatches 118; Indels 0; Gaps 0;
QY 98 agcgccagcgggaattgtacagagagactaccgcgcgaaggatagcccgacgaatagcaac 157
Db 1 rgrgrarararur 60
QY 158 caggaaacaaatctcatctgccaatgtccaccaggaacaaagaaagagactgagcagcc 217
Db 61 rsnrnrsnrnsnrnsnrnsnrnsnrnsnrnsnrnsnrnsnrnsnrnsnrnsnrnsnrn 120
QY 218 cctcgagaatgacaggaacacgcgcctgtggagaggtgagggccaccagcctgtgc 277
Db 121 rsnrnrsnrnsnrnsnrnsnrnsnrnsnrnsnrnsnrnsnrnsnrnsnrnsnrnsrn 180
QY 278 aaacaacaaacaaacaaacaaacataacacacacccacccagagggc 329
Db 181 rsnrnrsnrnsnrnsnrnsnrnsnrnsnrnsnrnsnrnsnrnsnrnsrgrgrgrgrc 232

RESULT 6
US-08-692-922-3/c
Sequence 3, Application US/08692922
Patent No. 6277592
GENERAL INFORMATION:
APPLICANT: Bidwell, Christopher A.
APPLICANT: Spurlock, Michael E.
TITLE OF INVENTION: PORCINE LEPTIN PROTEIN, NUCLEIC ACID
TITLE OF INVENTION: SEQUENCES CODING THEREFOR AND USES THEREOF
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: WHYTE HIRSCHBOECK DUDEK S.C.
STREET: Suite 2100 111 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: USA
ZIP: 53202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 31-JUL-1996
APPLICATION NUMBER: US/08/692,922
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Shaw, Melissa A.
REGISTRATION NUMBER: 38,301
REFERENCE/DOCKET NUMBER: PM-8935
TELECOMMUNICATION INFORMATION:
TELEPHONE: 414-273-2100
TELEFAX: 414-223-5000
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 435 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..435
US-08-692-922-3

Query Match
Best Local Similarity 5.0%; Score 34.8; DB 4; Length 435;
Matches 60; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 176 catggccaatgtccaccaggaacgaagagctggagcagccctgcagaatggacagga 235
Db 430 CAGGGCTGAGGTCCACGTCCGCGCATGTCTGTGACAGAGCCCTGTCAGCTGCTCAGGG 371
QY 236 acaccgcccgtgaggagggtgagggccaccagcctgtgc 277
Db 370 CCACCACCTCCGTGGAGTAGAGGGAGGCTTCCAGGACGCGC 329

RESULT 7
US-08-700-651-1
Sequence 1, Application US/08700651B
Patent No. 6015882
GENERAL INFORMATION:
APPLICANT: PETERSEN, CAROLYN
APPLICANT: LEECH, JAMES
APPLICANT: NELSON, RICHARD, C.
APPLICANT: GUT, JIRI
TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
TITLE OF INVENTION: INFECTIONS
FILE REFERENCE: 480.19-4(HV)
CURRENT APPLICATION NUMBER: US/08/700,651B
CURRENT FILING DATE: 1997-08-14
EARLIER APPLICATION NUMBER: 08/415,751
EARLIER FILING DATE: 1995-04-03
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 5163
TYPE: DNA
ORGANISM: Cryptosporidium parvum
US-08-700-651-1

Query Match
Best Local Similarity 5.0%; Score 34.8; DB 3; Length 5163;
Matches 39; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 279 aacaacaacaacaacaacaacataacacacacacacacacacacacacacacac 324
Db 874 aactacaacaacaacaacaacacacacacacacacacacacacacacacacac 919

RESULT 8
US-08-928-361B-4
Sequence 4, Application US/08928361B
Patent No. 6071518
GENERAL INFORMATION:
APPLICANT: Petersen, Carolyn
TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
TITLE OF INVENTION: SPECIES INFECTIONS
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETERS, VERNY, JONES & BIKSA
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-1840
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:


```
Query Match      8.5%; Score 59.8; DB 1; Length 7218;
Best Local Similarity 5.3%; Pred. No. 3.4e-08;
Matches 22; Conservative 227; Mismatches 164; Indels 0; Gaps 0;

QY 56 gccaggcagcgagcagattgactgaaagccagagatccaggcgagcggaattg 115
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1467 GCAAGTAGTAAGAGATAGACAATTCGTACRRRRRRRRRRRRRRRRRRRRR 1408

QY 116 acaggaggactacgcgcgaaggatagggccagaaatagcaaccaggaacaaatctcat 175
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1407 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1348

QY 176 catggccaatgtccaccaggaacaaagagctggagcagccctgcagatggacagga 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1347 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1288

QY 236 acaccgacctgtggagaggtgagggccaccagcctgctgcaaacacaaacaa 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1287 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1228

QY 296 ccacaaccataaacacacaccagcgaaggccaggtcgccgacctgcccctaaatt 355
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1227 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1168

QY 356 ccgatggccattcccaacaggcagatgaatgacgggttggtggagatggagatgat 415
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1167 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1108

QY 416 gaaatgttcaggaggagatagagagatccggagaaagcttagggagctac 468
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1107 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGCTCC 1055

RESULT 2
US-08-155-888-1
; Sequence 1, Application US/08155888
; Patent No. 6066623
; GENERAL INFORMATION:
; APPLICANT: Hoffman, Stephen L.
; APPLICANT: Hedstrom, Richard C.
; APPLICANT: Sedegah, Martha
; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE PROTECTIVE
; TITLE OF INVENTION: AGAINST MALARIA, METHODS OF PROTECTION AND VECTOR FOR
; TITLE OF INVENTION: DELIVERING POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Naval Medical Res. & Dev. Cmd.
; STREET: Bldg. 1, T-12 8901 Wisconsin Ave.
; CITY: Bethesda
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20889-5606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/155.888
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spevack, A. David
; REGISTRATION NUMBER: 24,743
; REFERENCE/DOCKET NUMBER: N.C. 75,851
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 295-6759
; TELEFAX: (202) 295-1022
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
```

```
; TOPOLOGY: circular
; HYPOTHEtical: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 1..755
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "promoter"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /label= CMV-IE
; OTHER INFORMATION: /note= "This feature acts as a promoter for any
; OTHER INFORMATION: downstream DNA sequence."
; OTHER INFORMATION: /citation= ([2])
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 933..2367
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /codon_start= 933
; OTHER INFORMATION: /function= "protein protective against malaria"
; OTHER INFORMATION: /product= "protein"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /number= 1
; OTHER INFORMATION: /label= IL2-CSP
; OTHER INFORMATION: /citation= ([1])
; PUBLICATION INFORMATION:
; AUTHORS: Sedegah, Martha
; AUTHORS: Hoffman, Stephen L.
; TITLE: Vaccination with Plasmodium yoelii CS protein
; TITLE: plasmid DNA protects against malaria
; JOURNAL: Science
; PUBLICATION INFORMATION:
; AUTHORS: Cullen, Bryan R.
; TITLE: TRANS-ACTIVATION OF HUMAN IMMUNODEFICIENCY
; TITLE: VIRUS OCCURS VIA A BIMODAL MECHANISM
; JOURNAL: CELL
; VOLUME: 46
; DATE: 26 SEP-1986
; PAGES: 973-982
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 4732
US-08-155-888-1

Query Match      6.4%; Score 44.6; DB 3; Length 5552;
Best Local Similarity 49.5%; Pred. No. 0.0013;
Matches 143; Conservative 0; Mismatches 144; Indels 2; Gaps 1;

QY 97 caggcgagcgagggaattgacaggaggactacgcccgaaggatagggcccaagaaatagcaa 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1770 CAAGGCCAGGAGCACCACAGGGGCCAGGAGCACCACCAAGGCGCAGGAGCACCACAAGGA 1829

QY 157 ccagga--aacaaaatctcatatgcccgaatgtccaccaggaagaaagagagctggagca 214
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1830 CCAGGAGCACCACAGGGTCCAGGAGCACCACCAAGGAGCAGGAGCACCACAAGGAGCAGGA 1889

QY 215 gccctctgcagaatggacaggaacacgcgcctgtggaggagagtgaggcgccaccagctgc 274
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1890 GCACCACAGGTCCAGGAGCACCACAGGGGCCAGGAGCACCACCAAGGCGCAGGAGCAGGA 1949

QY 275 tgcaaacacaaacacacacacacacacacacacacacacacacacacacacacacacac 334
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1950 CAAGAACCACCCCAACACCCACCCCAACCAACCAACCAACCAACCAACCAACCAACCA 2009

QY 335 tcgcccagctgccccttaacttcgatgggccaattcccaacagggcagatg 383
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2010 CAACAGCCACCACACAGCCACCACCAACCAACCAACCAACCAACCAACCAACCAATG 2058

RESULT 3
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 15:36:56 ; Search time 61.41 seconds
(without alignments)
2561.576 Million cell updates/sec

Title: US-09-327-750D-28

Perfect score: 700

Sequence: 1 acgagcgtctggccagcagc.....ggggcttggtggtccagtgga 700

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	59.8	8.5	7218	1	US-08-232-463-14, Appl
2	44.6	6.4	5552	3	US-08-155-888-1
3	38.4	5.5	7218	1	US-08-232-463-14
4	36.8	5.3	289	4	US-09-007-005-17
5	36.8	5.3	289	4	US-09-244-796-17
6	34.8	5.0	435	4	US-08-692-922-3
7	34.8	5.0	5163	3	US-08-700-651-1
8	34.8	5.0	5163	3	US-08-928-361B-4
9	34.8	5.0	5318	3	US-08-700-651-2
10	34.8	5.0	5318	3	US-08-928-361B-3
c 11	34.8	5.0	5917	4	US-08-692-922-1
c 12	34.4	4.9	1247	4	US-09-178-115-110
c 13	34.4	4.9	1247	4	US-09-177-776-110
c 14	34.4	4.9	2501	3	US-08-787-739-58
c 15	34.4	4.9	2501	4	US-09-178-115-58
c 16	34.4	4.9	2501	4	US-09-177-776-58
c 17	34.4	4.9	3532	3	US-08-787-739-90
c 18	34.4	4.9	3532	4	US-09-178-115-90
c 19	34.4	4.9	3532	4	US-09-177-776-90
c 20	34.4	4.9	10898	2	US-08-481-658B-5
c 21	34.4	4.9	10898	2	US-08-477-504A-5
c 22	34.4	4.9	10898	2	US-08-486-756A-5
c 23	34.4	4.9	10898	2	US-08-485-862B-5
c 24	34.4	4.9	10898	3	US-08-787-739-5
c 25	34.4	4.9	10898	3	US-08-487-077A-5
c 26	34.4	4.9	10898	3	US-08-485-863A-5
c 27	34.4	4.9	10898	4	US-08-485-049D-5

c 28	34.4	4.9	10898	4	US-09-178-115-5	Sequence 5, Appli
c 29	34.4	4.9	10898	4	US-09-177-776-5	Sequence 5, Appli
c 30	34.2	4.9	449	4	US-08-688-908-3	Sequence 3, Appli
31	34	4.9	198	5	PCT-US95-10668-1	Sequence 1, Appli
32	34	4.9	198	5	PCT-US95-10668-2	Sequence 2, Appli
33	34	4.9	198	5	PCT-US95-10668-3	Sequence 3, Appli
34	34	4.9	198	5	PCT-US95-10668-4	Sequence 4, Appli
35	33.6	4.8	3066	4	US-09-086-912-1	Sequence 1, Appli
36	33.4	4.8	5511	3	US-08-928-361B-2	Sequence 2, Appli
37	33.4	4.8	7334	3	US-08-928-361B-1	Sequence 1, Appli
38	33.2	4.7	413	1	US-08-303-266-1	Sequence 1, Appli
39	33.2	4.7	413	3	US-08-862-881-1	Sequence 1, Appli
40	33.2	4.7	688	4	US-08-998-416-915	Sequence 915, App
c 41	33.2	4.7	4279	4	US-09-041-886-22	Sequence 22, Appl
42	33	4.7	2251	2	US-08-836-443-2	Sequence 2, Appli
43	33	4.7	2648	2	US-08-836-443-1	Sequence 1, Appli
c 44	32.8	4.7	475	2	US-08-623-906A-20	Sequence 20, Appl
45	32.8	4.7	495	5	PCT-US95-11199-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 3.8%; Score 34; DB 2; Length 2277;
Best Local Similarity 29.0%; Pred. No. 2.8;
Matches 54; Conservative 38; Mismatches 94; Indels 0; Gaps 0;

QY 572 gaattgtctgcgtatcccttatgggggagctctctaatcaccatgaccatcatgatgaatt 631
DB 2224 RNARYTGTGTTRTAYTGTTCNACNARYTGRTRAANCNGTTCNGTGTTRTTNCCYT 2165

QY 632 ttgccttatgccttgactcctgccatttatcatgatgattaaactactgtgattcccgctgtt 691
DB 2164 TNGCYTGTTCNCKNSWNACYTGTTCNSWNSWARYTGTGTTCYTTCYTTCATGRT 2105

QY 692 ttcttttcttgcatcttcctaatatgcctttactgatccgttttgctgtgaaccctatg 751
DB 2104 TDATYTGNGGYTTNGGYTTTNGGRTGNACNGGYTTNACYTTNCCYTTRTCNCKNARNC 2045

QY 752 ttattt 757
DB 2044 KDATYT 2039

Search completed: March 11, 2002, 16:34:16
Job time: 3440 sec

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/676,967
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UCB96-055
TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-676-967-2

Query Match 3.8%; Score 34; DB 1; Length 2277;

Best Local Similarity 29.0%; Pred. No. 2.8;
Matches 54; Conservative 38; Mismatches 94; Indels 0; Gaps 0;

QY 572 gaattgctgcgtatccttctatggggagcgtctcttaacaccatgacccatgatgaatt 631
DB 2224 RNARVTTYGTTRTAYTCNACNARYTGRTRAACNKGTYTCNGTYTTRTNCCT 2165
QY 632 ttgccttgcctgactcctccattatcatgagattacgtgattccgcctgtt 691
DB 2164 TNGCYTTYTNCKNSWNACYGTGTCNSWNSNARYTGYTYTCYTCYTTCCAYTGR 2105
QY 692 ttcttttctgcattctcctaaatgcctttactgacccgtttgctgtaaccctatg 751
DB 2104 TDATVTGNGGYTTNGGYTTTNGGRTGNACNGYTTNACYTTNCCYTTTRTCNCKNARNC 2045
QY 752 ttattt 757
DB 2044 KDATYT 2039

RESULT 14
US-08-676-974-2/c
Sequence 2, Application US/08676974
Patent No. 5770422
GENERAL INFORMATION:
APPLICANT: COLLINS, KATHLEEN
TITLE OF INVENTION: Human Telomerase
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Science & Technology Law Group
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/676,974
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UCB96-055
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-676-974-2

Query Match 3.8%; Score 34; DB 1; Length 2277;

Best Local Similarity 29.0%; Pred. No. 2.8;
Matches 54; Conservative 38; Mismatches 94; Indels 0; Gaps 0;

QY 572 gaattgctgcgtatccttctatggggagcgtctcttaacaccatgacccatgatgaatt 631
DB 2224 RNARVTTYGTTRTAYTCNACNARYTGRTRAACNKGTYTCNGTYTTRTNCCT 2165
QY 632 ttgccttgcctgactcctccattatcatgagattacgtgattccgcctgtt 691
DB 2164 TNGCYTTYTNCKNSWNACYGTGTCNSWNSNARYTGYTYTCYTCYTTCCAYTGR 2105
QY 692 ttcttttctgcattctcctaaatgcctttactgacccgtttgctgtaaccctatg 751
DB 2104 TDATVTGNGGYTTNGGYTTTNGGRTGNACNGYTTNACYTTNCCYTTTRTCNCKNARNC 2045
QY 752 ttattt 757
DB 2044 KDATYT 2039

RESULT 15
US-09-098-487-2/c
Sequence 2, Application US/09098487
Patent No. 5917025
GENERAL INFORMATION:
APPLICANT: COLLINS, Kathleen
TITLE OF INVENTION: Human Telomerase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Science & Technology Law Group
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/098,487
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UCB96-055
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-098-487-2

Qy 74 gaaaaatggtgggttttgggggaaggaggttaaggggagaagcctggagggaggg 127
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 179 GAAAGGGGCGACGTCTCGGGGTTCGGGGTAAACGCACAAGTTCTCAGCACAGCG 126

; CURRENT APPLICATION NUMBER: US/09/244,796
 ;
 ; CURRENT FILING DATE: 1999-02-05
 ;
 ; EARLIER APPLICATION NUMBER: 60/035,963

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 40328 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-742-185-102

Query Match 5.0%; Score 44.4; DB 3; Length 40328;
Best Local Similarity 51.0%; Pred. No. 0.012;
Matches 105; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 684 cgcgtgtttcttcttcctgcaatttctcaatgcctttactgacgtttgtctgtga 743
DB 20178 CCCCTGCCCTTTTCTTCCATTGCTTGGTAGATCTTCTGTCATCCTTTTATTGGA 20237
QY 744 accctatgtattccatgtgtcaagtgggtctgtgttgccagcttctattgaagatt 803
DB 20238 GCCTATGTGTCTCTGACGTGAGATGGGTTTCTGTAATACAGCACATGATGGGTCTT 20297
QY 804 gctttgcactcagtgtaagtttctgtcagcagtagtttccaccatttgcattggaataa 863
DB 20298 GACTCTTATCCAGTTGCCAGCTGTGCTCTTTTAATTGGAGCATTTAGTCCATTGACAT 20357
QY 864 ttaagccaataaagcaatttaaaa 889
DB 20358 TTTAAAGTTAATATTGTTATGTATGAA 20383

RESULT 4
US-08-416-478A-1/c
; Sequence 1, Application US/08416478A
; Patent No. 5773578
; GENERAL INFORMATION:
; APPLICANT: Hercend, Thierry
; APPLICANT: Triebel, Frederic
; TITLE OF INVENTION: New Proteins Produced By Human
; TITLE OF INVENTION: Lymphocytes, DNA Sequences Encoding These Proteins And
; TITLE OF INVENTION: Their Pharmaceutical And Biological Uses
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 Seventh Street, N.W., Suite 400
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416.478A
; FILING DATE: 04-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/854,644
; FILING DATE: 08-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: YUN, Allen C.
; REGISTRATION NUMBER: 37,971
; REFERENCE/DOCKET NUMBER: HERCEND=1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1871 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE: CDS
; NAME/KEY: CDS
; LOCATION: 231..1724
US-08-416-478A-1

Query Match 4.2%; Score 37.2; DB 1; Length 1871;
Best Local Similarity 57.9%; Pred. No. 0.29;
Matches 66; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
QY 14 ctccataccgctctccattttgtgctgcaagctctgggaaagaatcccggaac 73
DB 239 CTCCACATCTCTCTATGCTGTGGCAGCGCTGAGCCCTCCAAAAGGAGGTTCAGAAAG 180
QY 74 gaaaaatggtggtttgggggaaggaggaagggagagaaagctggagggagg 127
DB 179 GAAAGGGGCGACTGGTGGGGGTGGGGGTAAAGGAGAGATTCTCCAGAGAGGG 126

RESULT 5
US-08-474-988B-1/c
; Sequence 1, Application US/08474988B
; Patent No. 5874250
; GENERAL INFORMATION:
; APPLICANT: Hercend, Thierry
; APPLICANT: Triebel, Frederic
; TITLE OF INVENTION: NEW PROTEINS PRODUCED BY HUMAN
; TITLE OF INVENTION: LYMPHOCYTES, DNA SEQUENCES ENCODING THESE PROTEINS AND
; TITLE OF INVENTION: THEIR PHARMACEUTICAL AND BIOLOGICAL USES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 Seventh Street, N.W., Suite 400
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474.988B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/416.478
; FILING DATE: 04-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/854,644
; FILING DATE: 08-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: YUN, Allen C.
; REGISTRATION NUMBER: 37,971
; REFERENCE/DOCKET NUMBER: HERCEND=1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1871 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE: CDS
; NAME/KEY: CDS
; LOCATION: 231..1724

```
Query Match          7.0%; Score 62.4; DB 1; Length 7218;
Best Local Similarity 4.5%; Pred. No. 2.4e-08;
Matches 18; Conservative 230; Mismatches 156; Indels 0; Gaps 0;

QY 52 ctgggaaagaaatcccggaagaaatggtggttgggggaaaggggtaaggggag 111
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1435 CRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1376

QY 112 aaagctggaggagggtttattgaggcccgtagagcgcgcgaactctaaag 171
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1375 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1316

QY 172 tgggaaaaaacaaataaaacacttgatatcagggtcctgaactcgtggtcaga 231
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1315 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1256

QY 232 gcaccaagcattcagctctctctcgttctgcttacttgcttcaagaacaaac 291
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1255 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1196

QY 292 cagaaaaaaaatctcatcgcaaatattcaccagaaacgaagagatggagcagc 351
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1195 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1136

QY 352 ctatgcagaatgagaggagacgcgccttgggagaggtgaagccaccagcctgcag 411
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1135 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1076

QY 412 gaaatgcagggagacaggtcgcgactgcgccttaatttcca 455
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1075 RRRRRRRRATCGCAAGTCCTCGACCTGACCGCAAGCTCGGA 1032

RESULT 2
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: ptzgt-F15
; US-08-232-463-14

Query Match          5.4%; Score 47.8; DB 1; Length 7218;
Best Local Similarity 3.9%; Pred. No. 0.00047;
Matches 10; Conservative 156; Mismatches 93; Indels 0; Gaps 0;

QY 593 ggggagcctctaaacacacatgacacatgatgaatttgcttgccttgactcct 652
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1052 GAGGAGCTGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1111

QY 653 gccatttatcatgagataactgtgattcccgctgtttctttcttcatttcc 712
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1112 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1171

QY 713 taatagccttactgacgttctgtggaacctatgtattccatgtgcaagtg 772
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1172 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1231

QY 773 gcttggttgccagctcttattgaagattgcttgcactcagtgtaagtctgtca 832
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1232 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1291

QY 833 gcagtagttccaccattt 851
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1292 YYYYYYYYYYYYYYYYYY 1310

RESULT 3
US-08-742-185-102
; Sequence 102, Application US/08742185
; Patent No. 6020476
; GENERAL INFORMATION:
; APPLICANT: Page, David C.
; APPLICANT: Reijo, Renee
; APPLICANT: Saxena, Richa
; APPLICANT: Hawkins, Trevor
; APPLICANT: Reeve, Mary Pat
; TITLE OF INVENTION: DAZ: A GENE FAMILY ASSOCIATED WITH AZOOSPERMIA
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,185
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/690,734
; FILING DATE: 31-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/310,429
; FILING DATE: 22-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WH194-07A2
```


THIS PAGE BLANK (USPTO)

/db_xref="taxon:9606"
/clone="NT2RP4000995"
/clone_lib="NT2RP4"
/cell_type="teratocarcinoma"
/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT 185 a 164 c 205 g 188 t 3 others
ORIGIN

Query Match 63.2%; Score 562.8; DB 10; Length 745;
Best Local Similarity 99.6%; Pred. No. 7e-107;
Matches 564; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 326 ccaggaacacgaagagatggaagcctctatgcagaaatggagaggaagcgcctttggg 385
|||||
Db 173 CCAGGAAACGAGAGATGGAGACGCTATGCAGAAATGGAGGAAGACCGCCCTTTGGG 232
|||||
QY 386 aggaagtgaagccaccagcctgcaggaatcgacgggagagcctcgcgacttgcgc 445
|||||
Db 233 AGAGGTGAAGCCACACGCTTCGGGGAATCGAGGGACAGGCTCGCGGACTTGCCTCC 292
|||||
QY 446 taatttcgatgggcccatacccaataggcagatcaatgatggatgggtggagatggaga 505
|||||
Db 293 TAATTTTCGATGGGCCATACCCCAATAGGCAGATCAATGATGGGATGGGATGGAGA 352
|||||
QY 506 tgatatgaataatctatgcagagagatgagagaaatcagaagaaacttagggagctgca 565
|||||
Db 353 TGATATGGAATATTCATGGAGGAGATGAGAGAAATCAGAGAAACTTAGGAGCTGCA 412
|||||
QY 566 gttgaggaattctgcgtatccttctatgggagcctcttaaccacatgacatcatga 625
|||||
Db 413 GTTGAGGAATGCTGCGTATCCTTATGGGGAGGCTCTTAATCACCATGACCATCATGA 472
|||||
QY 626 tgaatttgccttgccttgactcctgactcctgacttctatcatgagatgaataactgtattcc 685
|||||
Db 473 TGAATTTTGCCTTATGCTTGCCTGACTCCTGCCTATTCATGAGATTAATACTGTGATGCC 532
|||||
QY 686 gctgttttcttcttgcgtatcttcttaataatgacttactgactcgttgcgtgaac 745
|||||
Db 533 GCTGTTTTCTTTTCTTCTGCTATTTCTTAATATGCTTTTACTGATCCGTTTCTGTGAC 592
|||||
QY 746 cctatgtattccatgtgtcgaagtggtctgtgtgctcagcttctctattgaagatgc 805
|||||
Db 593 CTAATGTTATTCATGTCGTCGAGTGGGCTCTGTTGTTGTCGACCTTCTATTGAAGATTGC 652
|||||
QY 806 ctttgactcagtgtaagtctgtcagcagtagtttccaccatttgcagtgaaataatt 865
|||||
Db 653 CTTTGCACTCAGTGAAGTTTCTGTGACGAGTAGTTTCACCCATTTGCATGGCAAAATTT 712
|||||
QY 866 aaagccaataaagaactttaaaagc 891
|||||
Db 713 AAAGCTAATAAGCAATTTAAAGC 738
|||||

RESULT 14
AI598179/c
LOCUS AI598179 566 bp mRNA EST 12-MAY-1999
DEFINITION tn14e09.x1 NCI-CCAP_Brn25 Homo sapiens cDNA clone IMAGE:2167624 3', similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION AI598179
VERSION AI598179.1 GI:4607227
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 566)
AUTHORS NCI/NIHDS-CCAP et al.//www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project

(CGAP/BTGAP), Tumor Gene Index
Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 749 Std Error: 0.00
Seq Primer: 40UP from Gibco
High quality sequence stop: 453
POLYA-No.

FEATURES
source
1..566
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2167624"
/clone_lib="NCI-CCAP_Brn25"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCCGCATAGGTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 162 a 141 c 110 g 151 t 2 others
ORIGIN

Query Match 63.1%; Score 562.4; DB 10; Length 566;
Best Local Similarity 99.5%; Pred. No. 8.8e-107;
Matches 563; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 326 ccagaaaaacgaagatgagcagcctatgcagaaatgagaggaagcgcctttggg 385
|||||
Db 566 CCAGGANACGAGAGATGGAGCGCTATGCAGAAATGGAGGAAGACCGCCCTTTGGG 507
|||||
QY 386 aggaagtgaagccaccagcctgcaggaatcgacgggagagcctcgcgacttgcgc 445
|||||
Db 506 AGGAGTGAAGGCCACCGCTGCAGAAATCGACGGGACAGGCTCGCGACTTGCCTCC 447
|||||
QY 446 taatttcgatgggcccatacccaataggcagatcaatgatggatgggtggagatggaga 505
|||||
Db 446 TAATTTTCGATGGGCCATACNCAATAGGCAGATCAATGATGGGATGGGATGGAGA 387
|||||
QY 506 tgatatggaataattcagagagatgagaaatcagaagaaacttaggagctgca 565
|||||
Db 386 TGATATGGAATATTCATGAGAGATGAGAGAAATCAGAGAAACTTAGGAGCTGCA 327
|||||
QY 566 gttgaggaattctcgtcgtatccttatggggagctctctaatcaccatgaccatcatga 625
|||||
Db 326 GTTGAGGAATTTCTCGTATCTTATGGGAGGCTCTCTAATCACCATGACCATCATGA 267
|||||
QY 626 tgaatttgccttatgccttgactcctgcatttatcatgagattaactgtgattccc 685
|||||
Db 266 TGAATTTTGCCTTATGCTTGTACTCCTGCCATTTATCATGAGATTAATACTGTGATGCC 207
|||||
QY 686 gctgttttcttcttctcatttcttaataatgcctttactgactcgttctgtgtaac 745
|||||
Db 206 GCTGTTTTCTTTTCTTCTCATTTTCTTAATATGCTTTTACTGATCCGTTTGTGTGAAC 147
|||||
QY 746 cctatgtattccatgtgtcgaagtggtctgtgtgtgcagcttctctattgagattgc 805
|||||
Db 146 CCTATGTTATTCATGTCGAAGTGGGCTTGTGTGGCAGCTTCTATTGAAGATTGC 87
|||||

cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 759 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 451.

FEATURES

```

1. 574
Location/Qualifiers
Source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1640433"
/sex="male"
/lab.host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 162 a 141 c 111 g 160 t
ORIGIN
Query Match 63.3%; Score 564.4; DB 10; Length 574;
Best Local Similarity 99.8%; Pred. No. 3.4e-107;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 326 ccagagaaacgaagagatggagcagcctatgcagaatggagaggaagacgcgcctttggg 385
|||||
DB 574 CCAGAAAACGAAGAGATGGAGCAGCCTATGCAGAAATGGAGAGAGAGCGCCCTTTGGG 515
QY 386 aggaagtgaagccaccagcctgcaggaatcgacggagacagctcgccgacttgcgc 445
|||||
DB 514 AGAGGTGAAGCCACACGCTGCGAGGAATCGACGGGACAGGCTCGCCGACTTGCCTCC 455
QY 446 taatttcgatggggcaccatcccaataggcagatcaatgatggatgggtggagatggaga 505
|||||
DB 454 TAATTTTCATGGCGCATACCCCAATAGGCAGATCAATGATGGATGGGTGGAGATGGAGA 395
QY 506 tgatatggaatatcatcattgagagatgagagaaatcagagaaacacttagggagctgca 565
DB 394 TGATATGGAATATTTCATGTGAGGAGATGAGAGAAATCAGAGAAACTTAGGGAGCTGCA 335
QY 566 gttgaggaattgtctgcgtatccttatggggagctctctaatcaccatgaccatcatga 625
DB 334 GTTGAGGAATTCGTGCGTATCCTTATGGGGAGCTCTCTAATCACCATGACCATCATGA 275
QY 626 tgaatttgccttactgccttgccttgccttgccttgccttgccttgccttgccttgccttgc 685
DB 274 TGAATTTTGCCTTATGCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGC 215
QY 686 gctgttttcttttcttgccttgccttgccttgccttgccttgccttgccttgccttgccttgc 745
DB 214 GCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 155
QY 746 cctatgtatttccatgtgtcgaagtggtcttctgtgtgcagcttctctatttgaagatgc 805
DB 154 CTAATGTTATTTCCATGTGTCAGTGGGTCTTGTGTTGCCACCTCTCTATTGGAAGATGCG 95
QY 806 ctttgcactcagtgtaagtcttgcagcagtagtttccaccatttgcagtgagaaatatt 865
DB 94 CTTTGACATCAGTGAAGTTCTGTGCAGCAGTAGTTTCCACCATTTGCATGCAAAAATTT 35
QY 866 aaagccaataaagcaattttaaaaagc 891
|||||
DB 34 AAAGTAATAAAGCAATTTAAAAAGC 9

```

RESULT 11

```

AV709328 LOCUS 610 bp mRNA EST 09-OCT-2000
AV709328 DEFINITION ADC Homo sapiens cDNA clone ADCALB10 5', mRNA sequence.
AV709328 ACCESSION
AV709328 VERSION
AV709328.1 GI:10726849
EST.
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 610)
AUTHORS
Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao
H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,
Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu
G., Hu, R., Chen, J., Chen, Z., and Han, Z.
TITLE
Homo sapiens cDNA ADC clones
JOURNAL
Unpublished (2000)
COMMENT
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers
Source
1. 610
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ADCALB10"
/clone_lib="ADC"
/tissue_type="Adrenal gland"
/dev_stage="Adult"
/lab_host="SOLR"
/notes="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 181 a 115 c 148 g 166 t
ORIGIN
Query Match 63.3%; Score 564.4; DB 10; Length 610;
Best Local Similarity 99.8%; Pred. No. 3.4e-107;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 326 ccagagaaacgaagagatggagcagcctatgcagaatggagaggaagacgcgcctttggg 385
|||||
DB 22 CCAGAAAACGAAGAGATGGAGCAGCCTATGCAGAAATGGAGAGAGAGCGCCCTTTGGG 81
QY 386 aggaagtgaagccaccagcctgcaggaatcgacggagacagctcgccgacttgcgc 445
|||||
DB 82 AGAGGTGAAGCCACACGCTGCGAGGAATCGACGGGACAGGCTCGCCGACTTGCCTCC 141
QY 446 taatttcgatggggcaccatcccaataggcagatcaatgatggatgggtggagatggaga 505
DB 142 TAATTTTCATGGCGCATACCCCAATAGGCAGATCAATGATGGATGGGTGGAGATGGAGA 201
QY 506 tgatatggaatatcttgcgtatccttatggggagagctctctaatcaccatgaccatcatga 565
DB 202 TGATATGGAATATTTCATGTGAGGAGATGAGAGAAATCAGAGAAACTTAGGGAGCTGCA 261
QY 566 gttgaggaattgtctgcgtatccttatggggagagctctctaatcaccatgaccatcatga 625
DB 262 GTTGAGGAATTCGTGCGTATCCTTATGGGGAGGCTCTCTAATCACCATGACCATCATGA 321
QY 626 tgaatttgccttactgccttgccttgccttgccttgccttgccttgccttgccttgccttgc 685
DB 322 TGAATTTTGCCTTATGCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGC 381
QY 686 gctgttttcttttcttgccttgccttgccttgccttgccttgccttgccttgccttgccttgc 745
|||||

```


Email: est@watson.wustl.edu
This clone is available royalty-free through LNLN; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40RP from Gibco
High quality sequence stop: 437.

FEATURES
source

Location/Qualifiers
1. .662
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2519611"
/clone_lib="Schneider fetal brain 00004"
/sex="male"
/tissue_type="frontal lobe"
/dev_stage="5 months post-conception"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript SK (Stratagene);
Site_1: SstI; Site_2: XhoI; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence:
5'-GAGAGAGAGAGCTCAAGGATCCTTAATTAATTAATCCGCCCCCCC-3'
and 3' adaptor sequence:
5'-GAGAGAGAGCTCGAGCTTTT-3'. The library was
size-selected for >0.5 kb inserts and has an average
insert size estimated at 1.2 kb. This library was
constructed using the CAP-trapper method for full-length
enrichment and has not undergone amplification. Library
was constructed by Dr. Claudio Schneider (LNCIB-Area
Science Park, Trieste, Italy). 2 others
178 a 143 c 169 g 170 t

BASE COUNT
ORIGIN

Query Match 64.4%; Score 574; DB 10; Length 662;
Best Local Similarity 98.8%; Pred. No. 3.4e-109;
Matches 577; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 281 agaaaaaacacagaaaaaaatctcatcgcaaatattccaccaggaacgaaga 340
|||||
DB 65 AGAAAAACAACAGAAAAAATCTCATCGCAATATTCACCGAGAAACGAGCA 124
QY 341 gatgagagcctatgcagaaatggagaggaagacccctttgggagggtgaaggcca 400
|||||
DB 125 GATGGAGCAGCCTATGCAGAAATGGAGAGAACCCGCCCTTTGGGAGGAGTGAAGGCCA 184
QY 401 ccagcctgcagaaatcagcggggagcagctcgcaacttgccttaatttcgatgggc 460
|||||
DB 185 CCAGCCTGCAGGAATFCGACGGGGACAGGCTGCCGACTTGCCCTTAATTTTCGATGGGC 244
QY 461 cataccataaggcagatcaatgatgggatgggtggagatggagatgatggaataatt 520
|||||
DB 245 CATACCCATAGGCAGATCAATGATGGGATGGTGGGAGATGGAGATGATGGAATATT 304
QY 521 catggaggagatgagagaatcagaagaaacttagggagctgcagttgaggaattgtct 580
|||||
DB 305 CATGGAGGAGATCAGAGAAATCAGAGAAACTTAGGGAGCTGCAGTTCAGGAATTGTCT 364
QY 581 gcgtatccttatggggagctctctaataccatgacacatcatgatgaatttgccttat 640
|||||
DB 365 GCGTATCTTATGGGGAGCCTCTTAATCACCATGACCATCATGATGAATTTGCTTAT 424
QY 641 gcttgactcctgcattatcatgatgattaactgtgattccgcgtgtttcttttcc 700
|||||
DB 425 GCCTTGACTTCCTGCCATTTATCATGAGATTAATAGTGTATTCCTCTTTTNTC 484
QY 701 cttgcaatttcctaataatgccttactgatccggtttgctgtgaacctatgtattcca 760
|||||
DB 485 CTTGCAATTATCCTAATATGCCTTTACTGATCCGTTGCTGTGAACCTATGTTATTTCCA 544
QY 761 tgtgtcaagtgggtcttctgttgcagagcttctatttgaagattgcctttgactcaagtct 820
|||||
DB 545 TGTGTCAAGTGGGTCTTGTGTTGCCAGCTTCTATTITGAAGATTGTCCTNTGCACTCAAGTGT 604

QY 821 aagtttctgtcagcagtagtattccaccatttgcattgcatgaaaaatt 864
|||||
DB 605 AAGTTCGTCTCAGCAGTAGTTTTCACCCATCTGTCATGGAACATTT 648
RESULT 8
AA514324/c 603 bp mRNA 18-AUG-1997
LOCUS nf56b04.s1 NCI-CGAP.Co3 Homo sapiens cDNA clone IMAGE:923887 3'
DEFINITION similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74
(HUMAN);, mRNA sequence.
ACCESSION AA514324
VERSION AA514324.1 GI:2253832
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 603)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.
, Ph.D.
, cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbbrp/image/image.html
Insert Length: 840 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 433.
Location/Qualifiers
1. .603
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:923887"
/clone_lib="NCI-CGAP.Co3"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from 12 pooled bulk tumor samples and primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library went through one round of
normalization."
BASE COUNT 166 a 141 c 116 g 180 t
ORIGIN

FEATURES
source

Query Match 63.4%; Score 565.2; DB 10; Length 603;
Best Local Similarity 99.3%; Pred. No. 2.3e-107;
Matches 578; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 311 catggcaaatattccaccagg-aaaacgaagatggagcctatgcagaatggagagg 369
|||||
DB 603 CATGGCAATATATTCACGAGAAACGAGAGATGGAGACCCTATGCAATGGAGAGG 544
QY 370 aagaccgcctttgggaggagggtgaaggccaccagcctgcaggaatcgacggggacagg 429
|||||
DB 543 AAGACGCCCTTTGGGAGGAGGTGAAGCCACCAGCCTGCAGGAATCGACGGGACAGG 484
QY 430 ctgcgcgaattgccccctaatttttcgatggccatacccaatagcgagatcaatgatggga 489
|||||
DB 483 CTCGCCGACTTGCCCTTAATTTTCGATGGGCCATACCCAATAGGCAGATCAATGATGGGA 424

```

QY 672 atactgtattcccgctgtttctttcttgcatttccttaataatgctttactgacc 731
|||||
Db 226 ATACTGTGATTCCCGCTGTTTCTTTCTTCTTGCATTTTCTTAATATGCTTTACTGATC 167
|||||
QY 732 cgttgcgtgtaacctatgttattccatgctgtaagtgggtctctgttggccagcttc 791
|||||
Db 166 CGTTTGCTGTGAACCTATGTTATTTCCATGTCTCAAGTGGGTCTGTGTTGCCAGCTTC 107
|||||
QY 792 tattggaagattgccttgcctcactcagtgtaagtttctctgcacagtagtttcaaccattt 851
|||||
Db 106 TATTTGAAGATTGTCCTTTCACCTCAGTGTAAAGTTTCTGTGCAGCAGTAGTTTACCCATTT 47
|||||
QY 852 gcatggaataattaaagccaataaagcaatttaaaagc 891
|||||
Db 46 GCATGMAAAATTAAGCTAATAAGCAATTTAAAGC 7
|||||

RESULT 6
AI937016/c
LOCUS
DEFINITION
wp72b04.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2467279. 3'
similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74
(HUMAN);, mRNA sequence.
ACCESSION
AI937016
VERSION
AI937016.1 GI:5675886
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 606)
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(TITLE
(CGAP/BTGP), Tumor Gene Index
JOURNAL
Unpublished (1998)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 783 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 457.
FEATURES
Location/Qualifiers
1..606
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2467279"
/clone_lib="NCI_CGAP_Brn25"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not 1; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAACTGTAAGTGGAGCGCGCATAGCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT 168 a 144 c 119 g 173 t 2 others
ORIGIN

```

Query Match

64.8%; Score 577; DB 10; Length 606;

```

Best Local Similarity 97.9%; Pred. No. 8.2e-110;
Matches 594; Conservative 0; Mismatches 12; Indels 1; Gaps 1;
QY 285 aaacaacagaaaaaatactcatcgcaaatattaccaggaaacgaagagatg 344
|||||
Db 606 AAACACACAGAAAAAAATCTCATATGGCAATA-TCAACAGGANAACAGAGNATG 548
|||||
QY 345 gagcagcctatcgaaatgagagaaagaccgccctttgggagaggtgaaagccaccag 404
|||||
Db 547 GAGCAGCCTATCGAATGGAGAGGAGACCCGCCCTTTGGGAGGAGGTGAAGGCCACAG 488
|||||
QY 405 cctgcagaaaaatcgacggggacaggtcgccgacttgcccttaatttctcgatgggcata 464
|||||
Db 487 CCTGCAGCAAAATCGACGGGACAGGCTCGCCGACTTGCCTTAATTTTCATGGGCCATA 428
|||||
QY 465 cccaataggcagatcaatgattggatgggtgagatggagatgataatattatcatg 524
|||||
Db 427 CCCAATAGGCAGATCAATGATGGGATGGGTGGAGATGGAGATCATATGGAATATTTCATG 368
|||||
QY 525 gagagatgagaaaaatcagaagaaacttagggagctgagttgagggaattgtctgcgt 584
|||||
Db 367 GAGGAGATGAGAGAAATCAGAAGAAAACTTAGGGAGCTGCAGTTGAGGAATTTGCTGCCGT 308
|||||
QY 585 atccttatggggagctctcttaatacaccatgaccatcatgatgaattttgccttatgcct 644
|||||
Db 307 ATCCTTATGGGGAGCTCTCTTAATCACCATGACCATCATGATGAATTTTGCCTTATGCCT 248
|||||
QY 645 tgactcctgcatttatcatagagattaaactgatactgatactcccgctgtttcttcttcttctg 704
|||||
Db 247 TGACTCCTGCCAATTTATCATGAGATTAATPACTGTGATTCGCCGCTGTTTTCTTTTCTCTTG 188
|||||
QY 705 catttcttaataccttctactgacgttctgctgtaaacctatgttatttcttcttcttctg 764
|||||
Db 187 CATTTTCTTAATATGCCCTTTACTGATCCGTTTCTGCTGAACCTATGTTTATTTTCCATG 128
|||||
QY 765 tcaagtgggtctgtgtgctccagcttctatttgaagattgcttgcctcactcagtgtaagt 824
|||||
Db 127 TCAAGTGGGTCTGTGTGCCAGCTTCTATTGTAAGATTGCTTTGCCTTGCCTTGCCTTGCCTTGCCT 68
|||||
QY 825 ttctctcagcagtagtttcttccaccatttgcattggaataatccttaagcctaataagaattt 884
|||||
Db 67 TTCTCTCAGCAGTAGTTTTCACCCATTTGCATCGAAAAATTTAAAGCTAATAAAGCAATTT 8
|||||
QY 885 aaaaagc 891
|||||
Db 7 AAAAGC 1
|||||

RESULT 7
AI929106
LOCUS
DEFINITION
au5b10.y1 Schneider fetal brain 00004 Homo sapiens cDNA clone
IMAGE:2519611 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION
AI929106
VERSION
AI929106.1 GI:5665070
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 662)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin
J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
JOURNAL
Contact: Wilton RK
COMMENT
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

```

Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 339.

FEATURES
source

Location/Qualifiers
1. 620
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:107472"
/clone_lib="NCI_CGAP_Co9"
/tissue_type="colon tumor RER+"
/lab_host="DH10B"
/note="organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from RER+ colon tumor, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is not normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo (Soares4).

BASE COUNT 169 a 144 c 116 g 191 t
ORIGIN

Query Match 66.9%; Score 596.4; DB 10; Length 620;
Best Local Similarity 99.8%; Pred. No. 7.8e-114;
Matches 597; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 294 gaaaaaaatctctatcgcaaatattcaccaggaacgaagagatgagcagcct 353
|||||
Db 620 GAAAAAAAATCTCATCTGTCATATATTCACCGAGAAACGAGAGATGAGCAGCCT 561
|||||
QY 354 atcgagaatggagagaagaccgccccttgggagaggtgaaggccaccagcctcgagga 413
|||||
Db 560 ATGCAGATGGAGAGGAGAACCCCTTTGGGAGGAGGTGAAGGCCACCCAGCCTGCAGGA 501
|||||
QY 414 aatcgacggggacaggtcgccgacttgccttaatttccgatgggccaataccgaatagg 473
|||||
Db 500 AATCGACGGGGACAGGCTCGCCGACTTGCCTTAATTTTCGATGGGCCATACCCCAATAGG 441
|||||
QY 474 cagataatgatggatgggtgagatgagatgatgatgaaattcttccttgagagagatg 533
|||||
Db 440 CAGATCAATGATGGATGGGTGGAGATGAGATGATATGGAATATTCATGAGGAGATG 381
|||||
QY 534 agagaatcagaagaaaactaggagagtcgagtgagggaattgtcgtatcccttatg 593
|||||
Db 380 AGAGAAATCAGAAGAAAACCTTAGGGAGCTGCAGTTGAGGAATGTCTGCTATCCTTATG 321
|||||
QY 594 ggggagctctcaatcaccatgaccatgatgatgaatttgccttatgcttgccttgcctctg 653
|||||
Db 320 GGGGAGCTCTTAATCACCATGACCATCATGATGAATTTTGCTTATGCTTGCCTTGCCTCTG 261
|||||
QY 654 ccattttatcagatttaatactgtgattcccgctgttttcttcttcccttgcatcttct 713
|||||
Db 260 CCATTATCATGAGATTAACTATGATGATTCGCGCTGTTTTCTTTTCCCTTGCAATTTCTCT 201
|||||
QY 714 aatagcctttactgatccgtttgtgtgaaacctatgtttatttccatgtgcaagtggg 773
|||||
Db 200 AATATGCCCTTACTGATCCGTTTGTGCTGAACCCCTATGTTATTTCCATGTGCAAGTGGG 141
|||||
QY 774 tcttggttgcaggtctctatttgaagattgcttggactcagtgatgaagttctctgtag 833
|||||
Db 140 TCTTGTGTTGCCAGCTTCATTATGAAGATTGCTTTGCACTCAGTCAGTAAAGTTTCTGCTAG 81
|||||
QY 834 cagtagtttccaccatttgcattggaataattttaaagccaataaagcaattttaaagc 891
|||||
Db 80 CAGTAGTTTCACCCATTCGATGGAAAAATTTAAAGCTAATAAAGCAATTTAAAGAGC 23
|||||

RESULT 5
BF058872/c
LOCUS
DEFINITION 7k35g02.x1 NCI_CGAP_Ov18 Homo sapiens cDNA clone IMAGE:3477674 3' similar to SW:HG74_HUMAN Q00994 OVARIAN GRANULOSA CELL 13.0 KD EST 16-OCT-2000

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

PROTEIN HGR74. ; mRNA sequence.

BF058872
BF058872.1 GI:10812768

EST
human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 586)

AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL

Contact: Robert Strausberg, Ph.D.

COMMENT

Email: cgapbs@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento
Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The
I.M.A.G.E. Consortium DNA Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -400P from Gibco
High quality sequence stop: 464.
Location/Qualifiers
1. 586
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3477674"
/clone_lib="NCI_CGAP_Ov18"
/tissue_type="fibrotheoma"
/lab_host="DH10B (phage-resistant)"
/note="organ: ovary; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTCACCAATCTGAATGGAGCGCCGCCGACATTTTCTTTTCTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo.

FEATURES
source

BASE COUNT 166 a 143 c 112 g 164 t
ORIGIN

Query Match 64.8%; Score 577.4; DB 11; Length 586;
Best Local Similarity 99.7%; Pred. No. 6.8e-110;
Matches 578; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 312 atggcaaatattcaccaggaacgaagagatgagcagcctatgcagaatgagagaa 371
|||||
Db 586 ATGGCAAAATATTCACCGAGAAACGAAGAGATGAGCAGCAGTATGCAATGAGAGAGAA 527
|||||
QY 372 gaccgccccttgggagaggtgaagccaccagcctgcaggaaatcgacgggacaggct 431
|||||
Db 526 GACCGCCCTTTGGGAGGAGTGAAGCCACACAGCCTGCAGAAATCGACGGGACAGGCT 467
|||||
QY 432 cgccgacttgccttaatttgcgatgggccaatacccaataggcagatcaatgatggatg 491
|||||
Db 466 CGCCGACTTGCCTTAATTTTCATGGCCATACCCAAATAGGCAGATCAATGATGGGATG 407
|||||
QY 492 ggtggagatggagatgatgaaataatcatgagagatgagaaatcagaagaaaa 551
|||||
Db 406 GGTGGAGATGGAGATGATATGGAATATTCATGAGAGAGATGAGAGAAATCAGAAGAAAA 347
|||||
QY 552 cttagggagctcagttgaggaattgtctgctatctcttatggggagctctctaatcac 611
|||||
Db 346 CTTAGGGAGCTGCAGTTGAGGAATTCCTGCGTATCCTTATGGGAGCTCTCTAATCAC 287
|||||
QY 612 catgaccatcatgatgaatttgccttatgcttgccttgcctgcctgcattatcatgatgatta 671
|||||
Db 286 CATGACCATCATGATGAATTTTGCCTTATGCCCTTGACTCTCTGCCATTTATCATGAGATTA 227
|||||

/lab_host="DH10B"
/note="Organ: lymph node; Vector: pCMV-SPORT6; Site_1:
Salt; Site_2: NotI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size 1.25 kb. Life Technologies
catalog #: 11547-015"
BASE COUNT 169 a 145 c 120 g 178 t
ORIGIN

Query Match 68.2%; Score 607.8; DB 10; Length 612;
Best Local Similarity 99.7%; Pred. No. 3.4e-116;
Matches 609; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 281 agaaaaaacacagaaaaaaatctcatatgcaaaatattccaccggaagaaaga 340
|||||
Db 612 AGAAAAACCAACAGAAAAAATCTCATCATGCAAAATATTCACAGGAACGAAGA 553
QY 341 gatgagcagcctatgcagaatggaggaagacccctttggagaggtgaaggcca 400
|||||
Db 552 GATGAGCAGCCTATGCAAGATGGAGAGGAACCCCTTTGGGAGGAGGTGAAGGCCA 493
QY 401 ccagcctcaggaatcagcgggagcctcgccgacttgccttaatttcgatggc 460
|||||
Db 492 CCAGCCTCGAGGAATCGACGGGACAGCTCGCCGACTTGCCTTAATTTTCGATGGC 433
QY 461 catacccaatggcagatcaatgatggatggatggatggatggatggatggatgg 520
|||||
Db 432 CATACCAATAGGCAGATCAATGATGGATGGATGGATGGATGGATGGATGGAT 373
QY 521 catgagagatgagagaaatcagaagaaacttaggagctgcagtgaggaaattgt 580
|||||
Db 372 CATGGAGGAGATGAGAGAAATCAGAAAGAACTTAGGAGCTGCAGATTGTCT 313
QY 581 gcgtatcctatggggagctctcaatcacatgacatcatgatgaatttgccttat 640
|||||
Db 312 GCGTATCTTATGGGGAGCTCTTAATCACCATGACCATCATGATGAATTTGCCTTAT 253
QY 641 gcttgactcctgcctattatcatgagataactgattcccgctgtttcttttttc 700
|||||
Db 252 GCCTTGACTCTGCCATTTATCATGAGATTAATACTGTGATTCGCCGCTGTTCTTTTC 193
QY 701 ctgtatttcttaataatgcctttactgatccgtttgtgtgaacccctatgtattcca 760
|||||
Db 192 CTGTGATTTCTTAATATGCTTTTACTGATCGGTTGCTGTGAACCCCTATGTTTCCA 133
QY 761 tctgcaagtgggtctgtgtgcagctctctattgagattgcctttgcaactcagtg 820
|||||
Db 132 TGTGTCAAGTGGCTCTGTGTGGCAGCTTCTATTGTGAAGATTGCTTTGCACTCAGTGT 73
QY 821 aagtttctgcagcagtgatttccaccatttgcagtggaataatttaagccaataaagca 880
|||||
Db 72 AAGTTTCTGTACGAGTAGTTTACCCCATTTGCATGGAAAAATTTAAAGCTAATAAGCA 13
QY 881 atttaaaagc 891
|||||
Db 12 ATTTAAAAAGC 2

RESULT 2
AL193112/c
LOCUS
DEFINITION
qe69g09.x1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
IMAGE:1744288 3' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN); mRNA sequence.
ACCESSION
AL193112
VERSION
AL193112.1 GI:3744321
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 615)
AUTHORS
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE
JOURNAL
COMMENT
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 845 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 445.
Location/Qualifiers
1. 615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1744288"
/clone_lib="Soares_fetal_lung_NbHL19W"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: lung; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
15'-TGTACCAATCTGAAGTGGAGCGCGCAATTTTCTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M. Fatima Bonaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NBHL19W."
BASE COUNT 169 a 145 c 121 g 180 t
ORIGIN

Query Match 67.7%; Score 603; DB 10; Length 615;
Best Local Similarity 99.2%; Pred. No. 3.3e-115;
Matches 606; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 281 agaaaaaacacagaaaaaaatctcatatgcaaaatattccaccggaagaaaga 340
|||||
Db 615 AGAAAAACCAACAGAAAAAATCTCATGCGAAATATTCACAGGAACGAAGA 556
QY 341 gatgagcagcctatgcagaatggaggaagacccctttggagaggtgaaggcca 400
|||||
Db 555 GATGAGCAGCCTATGCAGATGGAGGAACGCCCTTTGGGAGGAGGTGAAGGCCA 496
QY 401 ccagcctcaggaatcagcgggagcagctcgccgacttgccttaatttcgatggc 460
|||||
Db 495 CCAGCCTCGAGAAATCGACGGGACAGCTCGCCGACTTGCCTTAATTTTCGATGGC 436
QY 461 catacccaatggcagatcaatgatggatggatggatggatggatggatggatgg 520
|||||
Db 435 CATACCAATAGGCAGATCAATGATGGATGGATGGATGGATGGATGGATGGAT 376
QY 521 catgagagatgagagaaatcagaagaaacttagggagctgcagttgagaaattgtc 580
|||||
Db 375 CATGGAGGAGATGAGAGAAATCAGAAAGAACTTAGGAGCTGCAGTTGTCT 316
QY 581 gcgtatcctatggggagcctctcaatcacatgacatcatgatgaatttgccttat 640
|||||
Db 315 GCGTATCTTATGGGGAGCTCTTAATCACCATGACCATGATGAATTTGCCTTAT 256
QY 641 gcttgactcctgcctattatcatgagataactgattcccgctgtttcttttttc 700
|||||
Db 255 GCCTTGACTCTGCCATTTATCATGAGATTAATACTGTGATTCGCCGCTGTTCTTTTC 196
QY 701 ctgtatttcttaataatgcctttactgatccgtttgtgtgaacccctatgtattcca 760
|||||
Db 195 CTGTCAATTTCTTAATATGCCCTTTACTGATCGTTTGTGTGAACCCATGATTAATTC 136
QY 761 tctgcaagtgggtctgtgtgccagctctctattgagattgcctttgcaactcagtg 820
|||||
Db 135 TGTGTCAAGTGGGTCTGTGTGGTGGCAGCTTCTATTGTGAAGATTGCTTTGCACACTCAGTGT 76

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 16:09:50 ; Search time 1609.39 Seconds
(without alignments)
5949.141 Million cell updates/sec

Title: US-09-327-750D-29
Perfect score: 891
Sequence: 1 accccatccccactctat.....aataaagcaatttaaaagc 891

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estfun:*

2: em_esthum:*

3: em_estin:*

4: em_estom:*

5: em_estpl:*

6: em_estba:*

7: em_estro:*

8: em_estov:*

9: em_htc:*

10: gb_estl:*

11: gb_est2:*

12: gb_htc:*

13: gb_gss:*

14: em_gss_fun:*

15: em_gss_hum:*

16: em_gss_inv:*

17: em_gss_pln:*

18: em_gss_pro:*

19: em_gss_rod:*

20: em_gss_vrt:*

21: em_gss_Other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	607.8	68.2	612	10 AW512400	AW512400 xx74b09.x
C 2	603	67.7	615	10 AI193112	AI193112 q669g09.x
C 3	597.2	67.0	626	10 AI928835	AI928835 au65b10.x
C 4	596.4	66.9	620	10 AA576958	AA576958 nm82d04.s
C 5	577.4	64.8	586	11 BF058872	BF058872 7k35g02.x
C 6	577.4	64.8	606	10 AI937016	AI937016 wp72b04.x
C 7	574	64.4	662	10 AI929106	AI929106 au65b10.y
C 8	565.2	63.4	603	10 AA514324	AA514324 nf56b04.s
C 9	564.4	63.3	568	10 AW131584	AW131584 xf31g01.x
C 10	564.4	63.3	574	10 AI075989	AI075989 ov47b05.x
C 11	564.4	63.3	610	10 AV709328	AV709328 av709328
C 12	562.8	63.2	574	10 BE222704	BE222704 hu51f06.x

13	562.8	63.2	745	10	AU132971	AU132971	AU132971	AU132971
C 14	562.4	63.1	566	10	AI598179	AI598179	AI598179	AI598179
C 15	560.8	62.9	566	11	BF732486	BF732486	BF732486	BF732486
C 16	560.8	62.9	720	10	AV701506	AV701506	AV701506	AV701506
C 17	560.4	62.9	566	10	AI859419	AI859419	AI859419	AI859419
C 18	559.8	62.8	584	10	AI188510	AI188510	AI188510	AI188510
C 19	559.6	62.8	837	10	BE728687	BE728687	BE728687	BE728687
C 20	557.8	62.6	578	10	AW156951	AW156951	AW156951	AW156951
C 21	557.8	62.6	599	10	AI033984	AI033984	AI033984	AI033984
C 22	555.2	62.3	606	10	AW088350	AW088350	AW088350	AW088350
C 23	555.2	62.3	779	11	BF764807	BF764807	BF764807	BF764807
C 24	553.6	62.1	676	11	BE896163	BE896163	BE896163	BE896163
C 25	553.4	62.1	716	11	BF716509	BF716509	BF716509	BF716509
C 26	552.4	62.0	573	11	BF940996	BF940996	BF940996	BF940996
C 27	551.8	61.9	772	11	BF339921	BF339921	BF339921	BF339921
C 28	550.8	61.8	858	11	BE793692	BE793692	BE793692	BE793692
C 29	550.4	61.8	690	11	BF825224	BF825224	BF825224	BF825224
C 30	549.8	61.7	589	10	AI188575	AI188575	AI188575	AI188575
C 31	549.6	61.7	556	10	AA573181	AA573181	AA573181	AA573181
C 32	549.6	61.7	728	11	BF972848	BF972848	BF972848	BF972848
C 33	549.4	61.7	566	10	AI033728	AI033728	AI033728	AI033728
C 34	548.8	61.6	736	10	BE408639	BE408639	BE408639	BE408639
C 35	548.8	61.6	879	11	BF338059	BF338059	BF338059	BF338059
C 36	548.6	61.6	701	11	BF403499	BF403499	BF403499	BF403499
C 37	548	61.5	674	11	BF745809	BF745809	BF745809	BF745809
C 38	547.6	61.5	775	11	BF706329	BF706329	BF706329	BF706329
C 39	546	61.3	707	10	BE732777	BE732777	BE732777	BE732777
C 40	545.4	61.2	704	11	BF488926	BF488926	BF488926	BF488926
C 41	543	60.9	570	10	AI887790	AI887790	AI887790	AI887790
C 42	540.4	60.7	730	11	BE900237	BE900237	BE900237	BE900237
C 43	538.2	60.4	621	10	AI570521	AI570521	AI570521	AI570521
C 44	537.4	60.3	574	10	AA843808	AA843808	AA843808	AA843808
C 45	536.6	60.2	682	10	BE727430	BE727430	BE727430	BE727430

ALIGNMENTS

RESULT 1
AW512400/c

LOCUS

DEFINITION

xx74b09.x1 NCI_CGAP_Lym12 Homo sapiens

similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74

(HUMAN); mRNA sequence.

AW512400

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

612 bp mRNA EST
03-MAR-2000
cdna clone IMAGE:2849369 3'
OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74

AW512400
xx74b09.x1 NCI_CGAP_Lym12 Homo sapiens
similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74
(HUMAN); mRNA sequence.

AW512400.1 GI:7150478
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 612)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

DNA Sequencing by: Washington University Genome Sequencing Center
Life Technologies catalog #: 11547-015
Cloned through the I.M.A.G.E. Consortium/LLNL at:
Image.llnl.gov/image/html/iresources.shtml
Seq primer: -40up from Gibco
High quality sequence stop: 422.

Location/Qualifiers
1..612
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2849369"
/clone_lib="NCI_CGAP_Lym12"
/tissue_type="lymphoma, follicular mixed small and large cell"

FEATURES
source
1..612
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2849369"
/clone_lib="NCI_CGAP_Lym12"
/tissue_type="lymphoma, follicular mixed small and large cell"


```
alignment_block:
US-09-327-750D-30 x AAF21748
Align seg 1/1 to: AAF21748 from: 1 to: 917
18 GluLysGluGluLysGluGluLysProGlnAspThrIleArgArgG1 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
306 CAGGAAACCAAGAGATGGAGCGCTATCCAGAATCGAGAGGAGACCG 355
34 uProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArgG 51
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
356 CCCTTTG.....GGAGAGGTGAAGGCCACCGCT...G 387
51 LyGlyArgArgArgPheArgValArgGlnProIleAlaHisTyrArgTIP 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
388 CAGGAATCAGCGGGACGCTGCGGACTTGCCTTAATTTTCGATGG 437
68 AspLeuMetGlnArg...ValGlyGluProGlnGlyArgMetArgGluG1 83
|||:|||||:|||||:|||||:|||||:|||||:|||||:
438 GCCATACCAATATTCATGGAGGAGATGAGAGAAATCAGAAGAAACTTA 534
83 uAsnValGlnArgPheGlyCysValArgGlnLeuMetGluLysLeuA 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
485 TGATATGGAATATTCATGGAGGAGATGAGAGAAATCAGAAGAAACTTA 534
100 rGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAsp...Pro 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
535 GGAGCTGCAGTTCAGGAATGTCTGCGTATCCTTTATGGGGGAGCTCT 584
116 ProHisAspHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
585 AATCACCATCACCATCATGATGAATTTGCTTATGCTTATGCT 623

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAC01005
seq_documentation_block:
AAC01005 standard; cDNA; 532 BP.
AC AAC01005;
XX
XX
XX 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 1003.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX P-PSDB; AAG00999.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 1003; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified within the
```


50 ArgGlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
 |||::: ||||| |||||::: ::::: |||
 329 GGAGGAATATCAGCGGGCGAGTTAGCGACTTGTCCCTAATTTTCG 378
 |||::: ||||| |||||::: ::::: |||
 66 gTrpAspLeuMetGlnArg...ValGlyGluProGlnGlyArgMetArg 82
 |||| ::::: |||| ::::: ||||
 379 ATGGGCCATACCTAATAGGCATATTGAGCACAATGAAGCGAGA..... 421
 |||| ::::: |||| ::::: ||||
 82 luGluAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLys 98
 ::::: ||||| ::::: ||||| ::::: |||||
 422 ..GATGATGTAGAAAGGTTTGTAGGCAGCATGCGAAATCAAGAGAAAG 469
 ||||| ::::: ||||| ::::: |||||
 99 LeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPr 115
 ||||| ::::: ||||| ::::: |||||
 470 ACTAGGGAACAGCAGATGAGGCACTATATGCGCTTCCAACTCCTGAACC 519
 ||||| ::::: ||||| ::::: |||||
 115 oProHisHisAspHisAspGluPheCysLeuMetPro 128
 | ||||| ::::: ||||| ::::: |||||
 520 T.....GACAACCATATGACTTTTGCCTCATACCT 550
 ||||| ::::: ||||| ::::: |||||

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT: AAC85548

seq_documentation_block:

ID AAC85548 standard; cDNA: 1364 BP.

AC AAC85548;

DT 04-JUN-2001 (first entry)

DE cDNA encoding CDIFF-4, Incyte ID No. 1990956CB1.

KW Human; cell differentiation; CDIFF; agonist; antagonist; epilepsy;
 cell proliferation; Alzheimer's disease; schizophrenia disorder;
 arteriosclerosis; cancer; atherosclerosis; diabetes mellitus; ss.

OS Homo sapiens.

FH Key Location/Qualifiers
 FT CDS 250..612
 FT /*tag= a
 FT /product= "CDIFF-4"

PN WO200119860-A2.

PD 22-MAR-2001.

PF 14-SEP-2000; 2000WO-US25435.

PR 15-SEP-1999; 99US-0154140.

PR 06-DEC-1999; 99US-0169155.

XX (INCY-) INCYTE GENOMICS INC.

PI Tang YT, Hillman JL, Yue H, Reddy R, Lal P, Shah P, Azimzai Y;
 PI Baughn MR, Lu DAM, Bandman O, Shih LL, Patterson C;

DR WPI; 2001-211447/21.

DR P-PSDB; AAB47126.

XX Isolated polypeptides and polynucleotides involved in cell
 PT differentiation are used for treatment, prevention and diagnosis of
 PT cell proliferative, developmental and neurological disorders e.g.
 PT cancer and Alzheimer's disease -

XX Claim 5; Page 121; 137pp; English.

PS The sequences given in AAC85545-72 encode human polypeptides involved
 CC in cell differentiation (CDIFF). CDIFF polypeptides and agonists of
 CC these are used to treat a disease or condition associated with
 CC decreased expression of functional CDIFF. An antagonist of CDIFF is
 CC used to treat a disease or condition associated with over expression
 CC of functional CDIFF. CDIFF polypeptides may be used for the treatment,

CC prevention and diagnosis of cell proliferative, developmental and
 CC neurological disorders, such as Alzheimer's disease, schizophrenic
 CC disorders, arteriosclerosis, cancer, atherosclerosis, diabetes mellitus
 CC and epilepsy. The CDIFF-4 sequence is homologous to Mus musculus
 CC REX-3. This sequence maps to chromosome 1 within the interval from
 CC 152.2 to 157.4 centimorgans, to chromosome 3 within the interval from
 CC 157.4 to 158.0 centimorgans, and to the X chromosome within the interval
 CC from 104.9 to 150.3 centimorgans.

SQ Sequence 1364 BP; 411 A; 269 C; 322 G; 362 T; 0 other;

alignment_scores:

Quality: 170.00 Length: 130
 Ratio: 1.954 Gaps: 6
 Percent Similarity: 66.923 Percent Identity: 34.615

alignment_block:

US-09-327-750D-30 x AAC85548 ..

Align seg 1/1 to: AAC85548 from: 1 to: 1364

1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
 ||||| ||||| ::::: ||||| |||||
 250 ATGGAGTCCAAAGAGAGAACTAGCGGCAAAACAATCTCAACGGGGAATTC 299
 ||||| ||||| ::::: ||||| |||||
 16 phiSGLnLysLysGluGluLysGluGluLysProGlnAspThrIleArgA 33
 ::::: ||||| ::::: ||||| ::::: |||||
 300 CCAACAAGAAACGAAGGAGGAGGAGCGGCCCC.....ACGCAGAAATG 343
 ||||| ::::: ||||| ::::: |||||
 33 rgGluProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaPro 49
 ::::: ||||| ::::: ||||| ::::: |||||
 344 AGAGAATCCCGCCCATTTG.....GGAGGGGTGAAGGCCAGAACCT 387
 ||||| ||||| ::::: ||||| |||||
 50 ArgGlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
 ||||| ||||| ||||| ||||| ::::: |||||
 388 GGAGGAATATCAGCGGGGGCGAGTTAGCGGACTTGTCCCTAATTTTCG 437
 ||||| ||||| ::::: ||||| |||||
 66 gTrpAspLeuMetGlnArg...ValGlyGluProGlnGlyArgMetArg 82
 |||| ::::: |||| ::::: ||||
 438 ATGGGCCATACCTAATAGGCATATTGAGCACAATGAAGCGAGA..... 480
 ||||| ||||| ::::: ||||| |||||
 82 luGluAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLys 98
 ::::: ||||| ::::: ||||| ::::: |||||
 481 ..GATGATGTAGAAAGGTTTGTAGGCAGCATGATGGAATCAAGAGAAAG 528
 ||||| ||||| ::::: ||||| |||||
 99 LeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPr 115
 ||||| ||||| ::::: ||||| |||||
 529 ACTAGGGAACAGCAGATGAGGCACTATATGCGCTTCCAACTCCTGAACC 578
 ||||| ||||| ::::: ||||| |||||
 115 oProHisHisAspHisAspGluPheCysLeuMetPro 128
 | ||||| ::::: ||||| ::::: |||||
 579 T.....GACAACCATATGACTTTTGCCTCATACCT 609
 ||||| ::::: ||||| ::::: |||||

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT: AAC23528

seq_documentation_block:

ID AAF23528 standard; DNA; 700 BP.

XX AAF23528;

AC AAF23528;

DT 22-MAR-2001 (first entry)

XX Mouse NADE DNA.

XX Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
 KW neurogenetic disease; NF-kappaB; ds.

XX Mus sp.

XX WO200075278-A2.

PD 14-DEC-2000.

THIS PAGE BLANK (USPTO)

```

      Quality: 81.00      Length: 121
      Ratio: 1.397
Percent Similarity: 47.934      Percent Identity: 23.967
      Gaps: 4
alignment_block:
US-09-327-750D-30 x US-08-459-448A-6
Align seg 1/1 to: US-08-459-448A-6 from: 1 to: 3624

14  GluAspHisGlnLysGluGluLysGluGluLysProGlnAspTh 30
||||:||||| :||||:||||| :||||:|||||
1026 GAACACGACGACCGCGGCCACCAACACGACATCAACCGGTGACCC 1075
30  rIleArgArgGluProAlaValAlaLeuIleSerGluAlaGlyLysAsnC 47
||||:||||| :||| :||||:
1076 TGGCGCTTCGCCGCGCGGACGTGTACCGCACCGAGAGCTACGCCGCGGTG 1125
47  ysAla.....ProArgGlyArgArgArg 55
:||||| :||||| :||||:
1126 CTGCTGTGGGGCATCTACCTGGAGCCCATCCAGCGCGTGCACCGGTGG 1175
56  PheArgValArgGlnProIleAlaHisTyrArgTrpAspLeuMetGlnAr 72
||||:||||| :||| :|||
1176 CTTCAACTTCACCAACCCCGACGACATCAGCGA.....CCGCG 1213
72  gValGlyGluProGlnGlyArgMetArgGluGluAsnValGlnArgPheG 89
| :||| :||| :|||||
1214 GCACCCCAACTACACCGCAGCCCTACGAGAGCCC..... 1247
89  lyGlyAspValArgGlnLeuMetGluLysLeuArgGluArg..... 102
||| :||| :||||:
1248 .....CGGCCTGCAGCTGAAGGACAGCGAGACCGGCTGCCCCC 1286
103 .....GlnLeuSerHisSerLeuArgAlaValSerThrAspProProH 117
:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
1287 CGAGACCGCGCGCGCCCACTACGAGAGCTACAGCCACCGCCTGAGCC 1336
117 sHisAspHisHis 121
||| :||| :|||
1337 ACATCGGCATCAT 1349
```


120 sHisAspGluPheCysLeuMetPro 128
 |||||
 596 TCATGATGAGTTTGGCNWATGCC 620

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF23529

seq_documentation_block:
 ID AAF23529 standard; DNA; 891 BP.

XX AAF23529;
 XX 22-MAR-2001 (first entry)
 XX Human NADE DNA.

XX Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
 KW neurogenetic disease; NF-kappaB; ds.
 XX Homo sapiens.

XX WO200075278-A2.
 PN 14-DEC-2000.
 PD 07-JUN-2000; 2000WO-US15621.

XX 07-JUN-1999; 99US-0327750.
 PR (UYCO) UNIV COLUMBIA NEW YORK.
 PA Sato T;

PI WPI: 2001-061707/07.
 DR New p75-neurotrophin receptor-associated cell death executor (NADE) and
 PT the gene encoding NADE, useful for modulating the activity of p75NTR
 PT and for detecting neurodegenerative diseases .
 XX Disclosure: Fig 1; 134pp; English.

CC The present invention relates to a purified polypeptide capable of
 CC binding neurotrophin receptor (p75-NTR). The invention is useful for
 CC binding and modulating the activity of p75NTR. The peptide mediates
 CC NGF-induced apoptosis, which plays an important role in neurogenetic
 CC diseases. The peptide of the invention and p75NTR are useful for
 CC inhibiting NF-kappaB activation in a cell or a subject, for inducing
 CC caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase
 CC and fragment nuclear DNA in a cell by co-expression of (I) and p75-NTR.
 XX Sequence 891 BP; 251 A; 182 C; 224 G; 234 T; 0 other;

alignment_scores:
 Quality: 171.50 Length: 129
 Ratio: 2.091 Gaps: 5
 Percent Similarity: 63.566 Percent Identity: 34.884

alignment_block:
 US-09-327-750D-30 x AAF23529 ..

Align seg 1/1 to: AAF23529 from: 1 to: 891

2 GluSerLysAspGlnGluValLysAsnLeuAsnMetGluAsnAspHisG1 18
 ::::|||||
 279 AAGAAAAAACACAGAAAAAAATCTCATGCGCAATATTCACCA 328
 18 nLysLysGluGluLysGluGluLysProGlnAspThrIleArgGluP 35
 ::::|||||
 329 GGAACAGAGATGAGATGAGCGCTATGCGAATGGAGAGACCGCC 378
 35 roAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArgGly 51
 || :::|

379 CTTTG.....GGAGGAGGTGAAGGCCACCGCCT...GCA 410
 52 GlyArgArgPheArgValArgGlnProIleAlaHisTyrArgTTPas 68
 |||||
 411 GGAATCGACGGGACAGCGCTCGCGACTTGCCTTAATTTTCGATGGCC 460
 68 pLeuMetGlnArg...ValGlyGluProGlnGlyArgMetArgGluGluA 84
 ::::|||||
 461 CATACCAATAGGACAGATCAATGATGGATGGGTGGA...GATGGAGATG 507
 84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
 ::::|||||
 508 ATATGGAATATTTCATGGAGGAGATGAGAAATCAGAAGAAACTTAGG 557
 101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAsp...ProPr 116
 |||||
 558 GAGCTGCAGTTGAGGAATTGCTGCGTATCTTATGGGGAGCTCTCTAA 607
 116 oHisHisAspHisAspGluPheCysLeuMetPro 128
 |||||
 608 TCACCATGACCATCATGATGAATTTGCCTTATGCCT 644

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517

seq_documentation_block:

ID AAH03517 standard; cDNA; 865 BP.

XX AC AAH03517;
 XX 26-JUN-2001 (first entry)

XX Human cDNA clone (5'-primer) SEQ ID NO:352.
 KW Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
 XX Homo sapiens.
 OS EP1074617-A2.
 XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.
 XX 29-JUL-1999; 99JP-0248036.
 XX 27-AUG-1999; 99JP-0300253.
 XX 11-JAN-2000; 2000JP-0118776.
 XX 02-MAY-2000; 2000JP-0183767.
 XX 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HE'Y RES INST.
 XX Ota T, Isoyui T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sogiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI: 2001-3*8749/34.
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs .
 XX Claim 1; SEQ ID 352; 2537pp + CD ROM; English.
 PS The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 865 BP; 239 A; 177 C; 224 G; 221 T; 4 other;

alignment_scores:
 Quality: 170.00 Length: 130
 Ratio: 1.954 Gaps: 6
 Percent Similarity: 66.923 Percent Identity: 34.615

alignment_block:

US-09-327-750D-30 x AAH03517 ..

Align seg 1/1 to: AAH03517 from: 1 to: 865

1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
 |||||
 191 ATGGAGTCCAAGAGAACTAGCGGCAACAATCTCAACGGGGAATGC 240
 16 pHISGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgA 33
 |||||
 241 CCAACAAGAAATCCCGCCATTG.....GGAGGGGTGAAGGCCAAGCCT 284
 33 rGluProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaPro 49
 |||||
 285 AAGAAGATCCCGCCATTG.....GGAGGGGTGAAGGCCAAGCCT 328
 50 ArgGlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
 |||||
 329 GGAGGAATATCAGCGGGCGAGTAGCGGACTTGTCCCTAATTTTCG 378
 66 gTRpAspLeuMetGlnArg...ValGlyGluProGlnGlyArgMetArgG 82
 |||||
 379 ATGGGCCATACCTAATAGGCATATTGACACAAATGAAGCGAGA..... 421
 82 LuGluAsnValGlnArgPheGlycylAspValArgGlnLeuMetGluLys 98
 |||||
 422 ..GATGATGTAGAAGGTTTGTAGGGCAGATGATGGAATCAAGAGAAG 469
 99 LeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspP 115
 |||||
 470 ACTAGGAACAGCAGATAGGCATATATGCGCTTCCAAACTCTGAACC 519
 115 oProHisHisAspHisAspGluPheCysLeuMetPro 128
 |||||
 520 T.....GACAACCATATGACTTTTGCCTCATACCT 550

seq_name: /SDS2/gcdata/geneseq/geneseq/NA2001.DAT:AAH13750

seq_documentation_block:

ID AAH13750 standard; cDNA; 1229 BP.

XX AAH13750;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:10656.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX

PN BP1074617-A2.

XX 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

DR Claim 8; SEQ ID 10656; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification

XX of the present invention.

SQ Sequence 1229 BP; 351 A; 243 C; 296 G; 339 T; 0 other;

alignment_scores:

Quality: 170.00 Length: 130
 Ratio: 1.954 Gaps: 6
 Percent Similarity: 66.923 Percent Identity: 34.615

alignment_block:

US-09-327-750D-30 x AAH13750 ..

Align seg 1/1 to: AAH13750 from: 1 to: 1229

1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
 |||||
 191 ATGGAGTCCAAGAGAACTAGCGGCAACAATCTCAACGGGGAATGC 240
 16 pHISGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgA 33
 |||||
 241 CCAACAAGAAACGAAGAGGGGAGGAGGCCGCC.....ACGCAAGATG 284
 33 rGluProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaPro 49
 |||||
 285 AAGAAGATCCCGCCATTG.....GGAGGGGTGAAGGCCAAGCCT 328


```
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR P-PSDB: AAM39425.
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
PT Claim 1; SEQ ID NO 784; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 862 BP; 235 A; 183 C; 238 G; 206 T; 0 other;
```

alignment_scores:

```
Quality: 450.00 Length: 129
Ratio: 3.879 Gaps: 2
Percent Similarity: 89.922 Percent Identity: 66.667
```

alignment_block:

US-09-327-750D-30 x AAI58581

Align seg 1/1 to: AAI58581 from: 1 to: 862

```
1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
226 ATGGAGTCCAAAGACAGACGCGTTAAACATCTCATCGTGGAAATGT 275
16 pHisGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgA 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
276 CAACACAGGAAATGATGAAAAAGATGAAAGGACGAAGTTGCTAATAAAG 325
33 rgGluProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaPro 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
326 GGGAGCCC...TTGGCCCTTACCTTGAATGTTAGTGAACTGTGTGCGCT 372
50 ArgGlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
373 ACAGGAAACCGTAGGCGGTTCCGGTTAGCAGCCCATCTCGCAGTATAG 422
66 gTrpAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluG 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
423 ATGGGACATAATGCATAGCTTGGAGAGCCACAGGCAAGGATGAGAGAG 472
83 luAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeu 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
473 AGAATATGGAAGAGATGGGGAGAGGTGAGAGCTGATGAAAGATG 522
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
523 ACGGAAACAGCATGTGAGTCATAGTTTGGCGGCGAGTCAGCACTGATCCCC 572
116 oHIsHisAspHisHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
573 TCACCATGACCATCATCATGATGAGTGTGCTTTGCTTATGCCC 609
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF59611
seq_documentation_block:
ID AAF59611 standard; cDNA; 898 BP.
XX AC AAF59611;
XX AC AAF59611;
XX 24-APR-2001 (first entry)
XX Human cell cycle and proliferation protein CCYPR-22 cDNA, SEQ ID NO:76.
XX Cell cycle and proliferation protein; CCYPR; human; agonist;
KW antagonist; gene therapy; detection; gene therapy;
KW transgenic animal disease model; immune disorder;
KW developmental disorder; cell signalling disorder;
KW cell proliferative disorder; cancer; tumour; andemia; epilepsy;
KW arteriosclerosis; asthma; allergy; diabetes mellitus;
KW menstrual cycle disorder; bacterial infection; ss.
XX Homo sapiens.
XX WO200107471-A2.
XX 01-FEB-2001.
XX 21-JUL-2000; 2000WO-US19948.
XX 21-JUL-1999; 99US-0145075.
XX 08-SEP-1999; 99US-0153129.
XX 10-NOV-1999; 99US-0164647.
XX (INCY-) INCYTE GENOMICS INC.
XX Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
PI Azimzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;
XX WPI: 2001-112727/12.
DR P-PSDB: AAB60474.
XX Human cell cycle and proliferation proteins and polynucleotides are
PT used to treat, diagnose and prevent immune, developmental and cell
PT signalling disorders and cell proliferative disorders including cancer -
XX Claim 5; Page 181-182; 205pp; English.
XX Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human
CC cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.
CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
CC associated with decreased expression of functional CCYPR, while CCYPR
CC antagonists are used to treat diseases or conditions associated with
CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
CC that specifically bind to CCYPR, and in drug screening methods to
CC identify compounds that modulate the activity of CCYPR. CCYPR
CC nucleotides can be used to generate transgenic animal models of human
CC disease, and can be used in gene therapy in target cells with genetic
CC abnormalities with respect to the expression of CCYPR for the
CC treatment or prevention of a disorder associated with CCYPR.
CC Diseases which can be diagnosed, treated and prevented using CCYPR
CC proteins, nucleic acids, agonists or antagonists include immune,
CC developmental and cell signalling disorders, and cell proliferative
CC disorders including cancer. Specific examples of these disorders
CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
CC diabetes mellitus, disorders of the menstrual cycle and infections
CC caused by bacteria.
XX Sequence 898 BP; 250 A; 186 C; 251 G; 211 T; 0 other;
```

```

alignment_scores:
  Quality: 450.00      Length: 129
  Ratio: 3.879         Gaps: 2
  Percent Similarity: 89.922  Percent Identity: 66.667

alignment_block:
US-09-327-750D-30 x AAF59611 ..
Align seg 1/1 to: AAF59611 from: 1 to: 898

1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
254 ATGGAGTCCAAAGAGGAGCAGCGTTTAAACAATCTCATCGTGGAAAATGT 303

16 pHisGlnLysGluGluLysGluGluLysProGlnAspThrIleArgA 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
304 CAACCGAGGAATAATGATGAAAAGATGAAAAGGAGCAAGTTGCTAATAAAG 353

33 rGluProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaPro 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
354 GGGAGGCC...TTGGCCCTACCTTTGAAATGTTAGTGAATACTGTGCCT 400

50 ArgGlyClyArgArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
401 AGAGGAACCGTAGCGGTTCCGCGTTAGGAGCCCATCTCGCAGTATAG 450

66 gTrpAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluG 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
451 ATGGGACATAATGATAGGCTTGGAGGCCACAGGCAAGGATGAGAGG 500

83 luAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeu 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
501 AGAATATGGAAGAGTATGGGAGGAGGTGAGACAGCTGATGGAAGCTG 550

100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
551 AGGGAAAGCAGTGTAGTCAATAGTCTCGGGCAGTCAGCACTGATCCCC 600

116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
601 TCACCATGACCATCATGATGAGTTTGGCCTTATGCCCC 637

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT :AAC03880

seq_documentation_block:
ID AAC03880 standard; cDNA; 662 BP.
XX
AC AAC03880;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 3878.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
DR P-PSDB; AAG03874.
XX

```

```

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 3878; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 662 BP; 177 A; 144 C; 192 G; 137 T; 12 other;

```

```

alignment_scores:
  Quality: 437.00      Length: 129
  Ratio: 3.902         Gaps: 2
  Percent Similarity: 86.822  Percent Identity: 65.116

alignment_block:
US-09-327-750D-30 x AAC03880 ..
Align seg 1/1 to: AAC03880 from: 1 to: 662

1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
207 ATGGAGTCCAAAGAGGAGCAGCGTTTAAACAATCTCATCGTGGAAAATGT 256

16 pHisGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgA 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
257 CAACGAGGAATAATGATGAAAAGATGAAAAGGAGCAAGTGWCTAATAAAG 306

33 rGluProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaPro 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
307 GGGAGGCC...TTGGCCCTACCTTTGRTATGTYTRGTGAATACTGTGCGCT 353

50 ArgGlyClyArgArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
354 AGAGGAATCGTAGCGGTTCCGCGTTAGGAGCCCATCTCGCAGTATAG 403

66 gTrpAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluG 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
404 ATGGGATATGATGCATAGGCTTGGAGAACCCACAGGCAAGGATGANAGAAG 453

83 luAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeu 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
454 AGAATATGGAAGAGTATGGGAGGAGGTGAGACAGCTGATGAAAAGCTG 503

100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
504 AGGGAAAGCAGTGTAGTCAATAGTCTCGGGCAGTCAGCACTGACCCCC 553

116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
554 TCACCATGACCATCATGATGAGTTTGCNNWATGCCCC 590

```

```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT :AAC03879

seq_documentation_block:
ID AAC03879 standard; cDNA; 698 BP.
XX
XX AAC03879;
XX

```

```

XX SQ Sequence 1229 BP; 351 A; 243 C; 296 G; 339 T; 0 other;
alignment_scores:
  Quality: 264.50      Length: 123
  Ratio: 3.112        Gaps: 4
  Percent Similarity: 69.106 Percent Identity: 47.967
alignment_block:
US-09-327-750D-12 x AAH13750
..
Align seg 1/1 to: AAH13750 from: 1 to: 1229
3 AsnValHisGlnGluAsnGluMetGluGln...ProLeuGlnAsnG1 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
236 AATGCCCAACAAGAAACAGAGGAGGGAGCGCCGCCACGCGAATGA 285
18 yGluGluAspArgProValGlyGlyGlyGluGlyHisGlnProAlaGlyA 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
286 AGAAGAAATCCCGCCATTGGGAGGGGTGAAGGCCAGAGCCTGGAGGAA 335
35 snAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisAsn 51
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
336 AT.....ATCAGCGCGGGG 349
52 GlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnArgG1 68
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
350 CGAGTAGGCGACTTCCCTAATTTTCGATGGCCATACCTAATAGGCA 399
68 nMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheMetG 85
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
400 TATTGACCAACAAGCG.....AGAGATGATGTAGAAAGTTTGTAG 443
85 luGluMetArgGluIleArgArgGlyLeuArgGluLeuGlnLeuArgAsn 101
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
444 GGCAGATGATGGAATCAAGAGAAGACTAGGCAACAGCAGATGAGGCAC 493
102 CysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisHisHisHis 118
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
494 TATATGCGCTTCCAACCTGCAACCTGACACCACTATTATGAC..... 535
118 pGluPheCysLeuMetPro 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
536 .....TTTTCCTCATACCT 550
seq_name: /SID52/gcgdata/geneseq/geneseq/NA2001.DAT:AAc85548
seq_documentation_block:
ID AAC85548 standard; cDNA; 1364 BP.
XX AC AAC85548;
XX DT 04-JUN-2001 (first entry)
XX DE cDNA encoding CDIFF-4, Incyte ID No. 1990956CB1.
XX KW Human; cell differentiation; CDIFF; agonist; antagonist; epilepsy;
KW cell proliferation; Alzheimer's disease; schizophrenia disorder;
KW arteriosclerosis; cancer; atherosclerosis; diabetes mellitus; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 250..612
XX FT /*tag= a
XX FT /product= "CDIFF-4"
XX PN WO200119860-A2.
XX PD 22-MAR-2001.
XX 14-SEP-2000; 2000WO-US25435.

```

```

XX 15-SEP-1999; 99US-0154140.
PR 06-DEC-1999; 99US-0169155.
XX (INCY-) INCYTE GENOMICS INC.
XX Tang YT, Hillman JL, Yue H, Reddy R, Lal P, Shah P, Azimzal Y;
PI Baughn MR, Lu DAM, Bandman O, Shih LL, Patterson C;
XX WPI; 2001-211447/21.
DR P-PSDB; AAB47126.
XX Isolated polypeptides and polynucleotides involved in cell
PT differentiation are used for treatment, prevention and diagnosis of
PT cell proliferative, developmental and neurological disorders e.g.
PT cancer and Alzheimer's disease -
XX Claim 5; Page 121; 137pp; English.
PS The sequences given in AAC85545-72 encode human polypeptides involved
XX in cell differentiation (CDIFF). CDIFF polypeptides and agonists of
CC these are used to treat a disease or condition associated with
CC decreased expression of functional CDIFF. An antagonist of CDIFF is
CC used to treat a disease or condition associated with over expression
CC of functional CDIFF. CDIFF polypeptides may be used for the treatment,
CC prevention and diagnosis of cell proliferative, developmental and
CC neurological disorders, such as Alzheimer's disease, schizophrenic
CC disorders, arteriosclerosis, cancer, atherosclerosis, diabetes mellitus
CC and epilepsy. The CDIFF-4 sequence is homologous to Mus musculus
CC REX-3. This sequence maps to chromosome 1 within the interval from
CC 152.2 to 157.4 centiMorgans, to chromosome 3 within the interval from
CC 157.4 to 158.0 centiMorgans, and to the X chromosome within the interval
CC from 104.9 to 150.3 centiMorgans.
XX SQ Sequence 1364 BP; 411 A; 269 C; 322 G; 362 T; 0 other;
alignment_scores:
  Quality: 264.50      Length: 123
  Ratio: 3.112        Gaps: 4
  Percent Similarity: 69.106 Percent Identity: 47.967
alignment_block:
US-09-327-750D-12 x AAC85548
..
Align seg 1/1 to: AAC85548 from: 1 to: 1364
3 AsnValHisGlnGluAsnGluMetGluGln...ProLeuGlnAsnG1 18
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
295 AATGCCCAACAAGAAACAGAGGAGGGAGCGCCGCCACGCGAATGA 344
18 yGluGluAspArgProValGlyGlyGlyGluGlyHisGlnProAlaGlyA 35
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
345 AGAAGAAATCCCGCCATTGGGAGGGGTGAAGGCCAGAGCCTGGAGGAA 394
35 snAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisAsnHis 51
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
395 AT.....ATCAGCGCGGGG 408
52 GlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnArgG1 68
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
409 CGAGTAGGCGACTTCCCTAATTTTCGATGGCCATACCTAATAGGCA 458
68 nMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheMetG 85
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
459 TATTGACCAACAAGCG.....AGAGATGATGTAGAAAGTTTGTAG 502
85 luGluMetArgGluIleArgArgGlyLeuArgGluLeuGlnLeuArgAsn 101
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
503 GGCAGATGATGGAATCAAGAGAAGACTAGGCAACAGCAGATGAGGCAC 552
102 CysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisHisHisHis 118
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



```

|||||
269 TGCAGAGGAAGACCGCCCTTTGGGAGGAGGTGAAGCCACCGCCTGCAG 318
|||||
34  lYAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisArgArg 50
|||||
319 GAAAT.....CGACGG 329
51  GlyGlnAlaArgArgLeuAlaProAsnPheArgTTPAlaIleProAsnAr 67
|||||
330 GGAMAGCTCGCGGAYTTGCCCTAATTTTCGATGGCCATACCCAATAG 379
|||||
67  gGlnMetAsnAspGlyLeuGlyAspGlyAspMetGluMetPheM 84
|||||
380 GCAGATCAATGATGGTGGCTGAGATGGATGATATGGAATATTC 429
84  etGluGluMetArgGluIleArgArgLysLeuArgGluLeuArg 100
|||||
430 TGGAGGAGATGAGAGAAATCAGAAGAACTTAGGGAGCTGCAKTTGAGG 479
101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisAspHisH 117
|||||
480 AATTGTCTGGGTATCTTATGGGGAKCTCTCTAATCACCATGACCATCA 529
117 s 117
530 T 530

```

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH01004

seq_documentation_block:

ID AAC01004 standard; cDNA; 485 BP.

AC AAC01004;

DT 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 1002.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping; ss.

OS Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

PA (GEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

DR P-PSDB; AAG00998.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for

PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for

PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 1002; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from

CC mRNAs encoding secreted proteins. An ORF has been identified within the

CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs

CC derived from 30 different tissues. EST sequences usually correspond

CC mainly to the 3' untranslated region (UTR) of the mRNA because they are

CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not

CC well suited for isolating cDNA sequences derived from the 5' ends of

CC mRNAs and even in those cases where longer cDNA sequences have been

CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from

CC mRNAs with intact 5' ends and can therefore be used to obtain full length

CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,

CC gene therapy and chromosome mapping procedures. They are used to obtain

CC upstream regulatory sequences and to design expression and secretion

XX vectors.

SQ Sequence 485 BP; 125 A; 113 C; 159 G; 84 T; 4 other;

alignment_scores:

Quality: 489.50 Length: 112

Ratio: 5.099 Gaps: 1

Percent Similarity: 85.714 Percent Identity: 81.250

alignment_block:

US-09-327-750D-12 x AAC01004 ..

Align seg 1/1 to: AAC01004 from: 1 to: 485

6 GlnGluAsnGluMetGluGlnProLeuGluAsnGlyGluGluAspAr 22

187 CAGAAACAGAGAGATGGAGACCCCTATGCAGATGGAGAGAACCG 236

22 gProValGlyGlyGlyGlnProAlaGlyAsnAsnAsnAsnA 39

237 CCCTTTGGGAGAGGTGAAGCCACCGCTGCAGGAAT..... 276

39 snAsnHisAsnHisAsnHisAsnHisArgArgGlyGlnAlaArgArg 55

277CGACGGGACAGGCTCGCCGA 297

56 LeuAlaProAsnPheArgTTPAlaIleProAsnArgGlnMetAsnAspG 72

298 CTTGCCCTAATTTTCGATGGCCATACCCAATAGGCAGATCAATGATGG 347

72 yLeuGlyGlyAspGlyAspMetGluMetPheMetGluGluMetArgG 89

348 GATGGGTGAGATGGAGATGATATGGAATATTCATGGAGAGATGAGAG 397

89 lulleArgArgLysLeuArgGluLeuGlnLeuArgAsnCysLeuArgIle 105

398 AAATCAGAGAGAACTTAGGGAGCTGCAKTTGAGGAATTTGTCGCTATC 447

106 LeuMetGlyGluLeuSerAsnHisAspHisHis 117

448 CTTATGGGGAKCTCTCTAATCACCATGACCATCAT 483

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517

seq_documentation_block:

ID AAH03517 standard; cDNA; 865 BP.

XX AC AAH03517;

XX 26-JUN-2001 (first entry)

XX Human cDNA clone (5'-primer) SEQ ID NO:352.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

```

XX WPI; 2000-611515/58.
DR P-PSDB; AAB58845.
XX
PT New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases.
XX
XX Claim 1; Page 581-582; 1299pp; English.
XX
CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC antidiabetic; antiinflammatory; antiviral; antiallergic; hepatotropic;
CC antibacterial; antifungal; antiparasitic and cardiant activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and agonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemias; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.
XX
SQ Sequence 917 BP; 228 A; 203 C; 270 G; 211 T; 5 other;

alignment_scores:
  Quality: 550.50      Length: 119
  Ratio: 5.193        Gaps: 1
  Percent Similarity: 89.076      Percent Identity: 84.874

alignment_block:
US-09-327-750D-12 x AAF21748 ..
Align seg 1/1 to: AAF21748 from: 1 to: 917
6 GlnGluAsnGluGluMetGluGlnProLeuGlnAsnGlyGluGluAspar 22
306 CAGGAAACCAAGAGATGGAGCAGCCTATGCAGATGGAGAGAACCG 355
22 gProValGlyGlyGlyGlyHisGlnProAlaGlyAsnAsnAsnA 39
356 CCCTTTGGGAGAGGTGAAGGCCACCAGCCTGCAGGAAT..... 395
39 snAsnHisAsnHisAsnHisAsnHisHisArgArgGlyGlnAlaArg 55
396 .....CGACGGGCAGACGGCTGCCCGA 416
56 LeuAlaProAsnPheArgTrpAlaIleProAsnArgGlnMetAsnAsp 72
417 CTTGGCCCCCAATTTTCGATGGGCCATACCCCAATAGCAGATCAATG 466
72 yLeuGlyGlyAspGlyAspMetGluMetPheMetGluGluMetArg 89
467 GATGGGTGGAGATGGAGATGATATGGAATATTCATGGAGGAGATGAG 516
89 luIleArgArgLysLeuArgGluLeuGlnLeuArgAsnCysLeuArg 105
517 AAATCAGAAAGAACTTAGGAGCTCGATGAGGAATTCCTGCGCATC 566
106 LeuMetGlyGluLeuSerAsnHisHisAspHisHisAspGluPheCys 122
567 CTTATGGGGGAGCTCTCTAATCACCATGACCATCATCATGATGAAT 616
122 uMetPro 124

```

```

|||||||
617 TATGCCT 623

seq_name: /SID62/gcgdata/geneseq/geneseq/NA2000.DAT:AAC01005
seq_documentation_block:
ID AAC01005 standard; cDNA: 532 BP.
XX
AC AAC01005;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 1003.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (BEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
WPI; 2000-500381/45.
DR P-PSDB; AAG00999.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
PS Claim 1; SEQ ID 1003; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dr primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 532 BP; 151 A; 121 C; 163 G; 91 T; 6 other;

alignment_scores:
  Quality: 504.50      Length: 117
  Ratio: 5.096        Gaps: 1
  Percent Similarity: 84.615      Percent Identity: 79.487

alignment_block:
US-09-327-750D-12 x AAC01005 ..
Align seg 1/1 to: AAC01005 from: 1 to: 532
1 MetAlaAsnValHisGlnGluAsnGluGluMetGluGlnProLeuGlnAs 17
219 ATGGCAATATTATCACCAGGAAACGAAGATGGAGCAGCCTATGCAGAA 268
17 nGlyGluGluAspArgProValGlyGlyGlyGlyHisGlnProAlaG 34

```


OM of: US-09-327-750D-12 to: N_Geneseq_1101.* out_format : pfs
Date: Mar 11, 2002 3:42 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame+_B2n.model -DEV=xlh
-Q=/cgn2_1/USPIO_Spool/US09327750/runat_11032002_101154_20362/app_query.fasta_1.1472
-DB=N_Geneseq_1101 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-XGAPOP=4.500 -XGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-XGAPOP=6.000 -XGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09327750_@CGN1_1_330 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:
Query: US-09-327-750D-12
Query length: 124
Database: N_Geneseq_1101.*
Database sequences: 930621
Database length: 428662619
Search time (sec): 355.560000

score_list:	Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF23528 +	700	!	685.00	1330.85	4.7e-66	700
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAF23529 +	891	XX	576.50	1115.82	4.5e-54	891
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAF21748 +	917	!	550.50	1064.57	3.2e-51	917
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAAC01005 +	532	PT	504.50	979.69	1.7e-46	532
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAAC01004 +	485	!	489.50	951.18	6.6e-45	485
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517 +	264	!	264.50	504.49	5.0e-20	264
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13750 +	1229	XX	264.50	504.49	7.8e-20	1229
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH82386 +	1364	PS	264.50	500.07	8.9e-20	1364
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH87147 +	187	!	193.00	379.21	4.8e-13	187
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143 +	792	!	171.50	327.55	6.2e-11	792
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH75910 +	767	!	171.50	323.35	6.2e-10	767
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH58581 +	862	!	171.50	322.22	7.2e-10	862
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH59611 +	898	!	171.50	321.82	7.5e-10	898
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH82386 +	426	!	168.00	322.21	7.2e-10	426
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAAC03880 +	662	!	162.50	307.14	5.0e-09	662
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAAC03879 +	692	!	162.50	306.63	5.3e-09	692
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH60367 +	858	!	158.50	296.78	1.9e-08	858
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAH90866 +	2273	!	101.00	174.60	0.1195	2273
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH32130 +	2273	!	101.00	174.60	0.1195	2273
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAH60302 +	2812	!	101.00	172.54	0.1558	2812
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH47157 +	1330	!	99.50	177.85	0.0788	1330
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAAC01003 +	401	!	99.50	188.51	0.0201	401
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH49183 +	1066	!	95.50	171.17	0.1855	1066
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH38965 +	1165	!	95.50	170.31	0.2072	1165
/SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT:AAH91184 +	2270	!	95.50	163.83	0.4755	2270
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH49164 +	2982	!	95.50	161.18	0.6680	2982
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH75288 +	2740	!	94.50	160.05	0.7730	2740
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH29736 +	2660	!	92.50	156.41	1.23	2660
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH08446 +	745	!	92.00	167.79	0.2861	745
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH37255 +	407	!	90.00	169.75	0.2228	407
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH72409 +	1197	!	87.50	154.37	1.60	1197
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH95484 +	1968	!	86.50	147.58	3.83	1968
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH87111 +	1968	!	86.50	147.58	3.83	1968
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH06449 +	2405	!	86.50	145.63	4.91	2405
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH14082 +	397	!	85.00	160.19	0.7592	397
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH135463 +	397	!	85.00	160.19	0.7592	397
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH103935 +	397	!	85.00	160.19	0.7592	397
/SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT:AAH37948 +	2353	!	84.50	141.92	7.90	2353
/SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:AAH07530 +	2353	!	84.50	141.92	7.90	2353

/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH02898 +	84.50	141.92	7.90	2353
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH41352 -	84.00	157.43	1.08	431
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH70218 +	83.50	146.40	4.45	1212
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH81469 +	83.50	108.36	584.58	6087
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH21610 -	83.50	91.38	5.2e+03	3499
seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF23528				
seq_documentation_block:				
ID AAF23528 standard: DNA: 700 BP.				
AC AAF23528;				
DT 22-MAR-2001 (first entry)				
XX Mouse NADE DNA.				
DE Neurotrophin receptor; p75-NTR; NGF-Induced apoptosis;				
KW neurogenetic disease; NF-kappaB; ds.				
XX Mus sp.				
XX WO200075278-A2.				
XX 14-DEC-2000.				
XX 07-JUN-2000; 2000WO-US15621.				
XX 07-JUN-1999; 99US-0327750.				
XX (UYCO) UNIV COLUMBIA NEW YORK.				
XX Sato T;				
XX WPI; 2001-061707/07.				
XX New p75-neurotrophin receptor-associated cell death executor (NADE) and the gene encoding NADE, useful for modulating the activity of p75NTR and for detecting neurodegenerative diseases -				
XX Claim 12; Fig 1; 134pp; English.				
XX The present invention relates to a purified polypeptide capable of binding neurotrophin receptor (p75-NTR). The invention is useful for binding and modulating the activity of p75NTR. The peptide mediates NGF-induced apoptosis, which plays an important role in neurogenetic diseases. The peptide of the invention and p75NTR are useful for inhibiting NF-kappaB activation in a cell or a subject, for inducing caspase-2 and caspase-3 activity to cleave poly (Adp-ribose) polymerase and fragment nuclear DNA in a cell by co-expression of (1) and p75-NTR.				
XX Sequence 700 BP; 177 A; 188 C; 203 G; 132 T; 0 other;				
alignment_scores:				
Quality: 685.00 Length: 124				
Ratio: 5.524 Gaps: 0				
Percent Similarity: 100.000 Percent Identity: 97.581				
alignment_block:				
US-09-327-750D-12 x AAF23528				
Align seg 1/1 to: AAF23528 from: 1 to: 700				
1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17				
177 ATGCCCAATGTCACCAAGAAACGAAGAGCTGGAGCAGCCCTGCAGAA 226				
17 nGlyGluGluAspArgProValGlyGlyGlyHisGlnProIaG 34				
227 TGACAGGAAGACCGCCCTGTGGAGAGAGTGGAGCCACGACCTGCTG 276				
34 lyAsnAsnAsnAsnAsnAsnHisHisAsnHisHisHisArgArg 50				


```

; TELEFAX: (202) 295-1022
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 5552 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   FEATURE:
;     NAME/KEY: promoter
;     LOCATION: 1..755
;     IDENTIFICATION METHOD: experimental
;     OTHER INFORMATION: /function= "promoter"
;     OTHER INFORMATION: /evidence= EXPERIMENTAL
;     OTHER INFORMATION: /label= CMV-IE
;     OTHER INFORMATION: /note= "This feature acts as a promoter for any
;     OTHER INFORMATION: downstream DNA sequence."
;     OTHER INFORMATION: /citation= ([2])
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 933..2367
;     IDENTIFICATION METHOD: experimental
;     OTHER INFORMATION: /codon_start= 933
;     OTHER INFORMATION: /function= "protein protective against malaria"
;     OTHER INFORMATION: /product= "protein"
;     OTHER INFORMATION: /evidence= EXPERIMENTAL
;     OTHER INFORMATION: /number= 1
;     OTHER INFORMATION: /label= IL2-CSP
;     OTHER INFORMATION: /citation= ([1])
;   PUBLICATION INFORMATION:
;     AUTHORS: Sedegah, Martha
;     AUTHORS: Hedstrom, Richard C.
;     AUTHORS: Hoffman, Stephen L.
;     TITLE: Vaccination with Plasmodium yoelii CS protein
;     TITLE: plasmid DNA protects against malaria
;     JOURNAL: Science
;     PUBLICATION INFORMATION:
;     AUTHORS: Cullen, Bryan R.
;     TITLE: TRANS-ACTIVATION OF HUMAN IMMUNODEFICIENCY
;     TITLE: VIRUS OCCURS VIA A BIMODAL MECHANISM
;     JOURNAL: CELL
;     VOLUME: 46
;     PAGES: 973-982
;     DATE: 26 SEP-1986
;     RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 4732
;   US-08-155-888-1

```

```

alignment_scores:
  Quality: 78.50      Length: 35
  Ratio: 3.738       Gaps: 1
  Percent Similarity: 60.000   Percent Identity: 45.714

alignment_block:
  US-09-327-750D-12 x US-08-155-888-1 ..
  Align seg 1/1 to: US-08-155-888-1 from: 1 to: 5552
  12 GluGlnProLeuGlnAsnGlyGluGluAspArgProValGlyGlyGly 28
  ::: ||| ||| ::::: |||||
  2013 CAGCCACCAACACAGCCACCAACCAACCAACCAACCAACCAACCA 2048
  28 uGlyHisGlnProAlaGlyAsnAsnAsnAsnAsnHisAsnHisAsnH 45
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  2049 .....CAGCCAGAGGGTAATAACCAACCAACCAACCAACCAAT 2091
  45 isAsn 46
  ::|||
  2092 ATAAT 2096
seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-09-023-173-4

```

```

seq_documentation_block:
; Sequence 4, Application US/09023173
; Patent No. 6066781
; GENERAL INFORMATION:
;   APPLICANT: Sutliff, Thomas D.
;   APPLICANT: Rodriguez, Raymond L.
;   TITLE OF INVENTION: Production of Mature Proteins
;   TITLE OF INVENTION: in Plants
;   NUMBER OF SEQUENCES: 23
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Dehlinger & Associates
;     STREET: 350 Cambridge Ave., Suite 250
;     CITY: Palo Alto
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94306
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FastSeq for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/023,173
;     FILING DATE: 13-FEB-1998
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 60/038,168
;     FILING DATE: 13-FEB-1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Petithory, Joanne R
;       REGISTRATION NUMBER: P42995
;       REFERENCE/DOCKET NUMBER: 0665-0007.30
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 650-324-0880
;       TELEFAX: 650-324-0960
;     INFORMATION FOR SEQ ID NO: 4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 1140 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     IMMEDIATE SOURCE:
;       CLONE: codon-optimized Ramy3D signal-prosutilisin BPN'
;     US-09-023-173-4
alignment_scores:
  Quality: 77.00      Length: 123
  Ratio: 1.242       Gaps: 5
  Percent Similarity: 50.407   Percent Identity: 23.577

alignment_block:
  US-09-327-750D-12 x US-09-023-173-4 ..
  Align seg 1/1 to: US-09-023-173-4 from: 1 to: 1140
  5 HisGlnGluAsnGluMetGluGlnProLeuGlnAsnGly..... 18
  |||:::|||||:::|||||
  396 CATCGACAGCGGTCGACAGCAGCCACCGGACCTCAAGTCGCGCGG 445
  19 .....GluGluAspArgProValGlyGlyGlyGlyGlyH 30
  |||:::|||||:::|||||
  446 GAGCTAGCAGTGTCCCGGAGGAGCAGCAACCCGTT.....CCAGGAC 486
  30 isGlnProAlaGlyAsnAsnAsnAsnAsnHisAsnHisAsnHisAsn 46
  ||||| ::::: ||||| ||||| |||||
  487 ACCAACAGCCATGGCACCACCGTCGCGGACCGTCGCGCGCTCACCAA 536
  47 HisHisArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTpa 63
  ::||| ||||| ::|||
  537 CAGCATCGGCGTCTCGCGTCGC..... 560
  63 alleProAsnArgGlnMetAsnAspGlyLeuGlyGlyAspGlyAspAspM 80

```

```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 663..3164
;
US-08-396-001-3

alignment_scores:
    Quality: 79.00      Length: 92
    Ratio: 1.362      Gaps: 5
    Percent Similarity: 63.043      Percent Identity: 26.087

alignment_block:
US-09-327-750D-12 x US-08-396-001-3 ..
Align seg 1/1 to: US-08-396-001-3 from: 1 to: 3455

17 AsnGlyGluGluAspArgProValGlyGlyGlyGlyGlyGlyHisGlnProAl 33
||||:||||: :|||||
2451 AACACGAAATAAGAACCC.....CATAACAAAAA 2482

33 aGlyAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArgA 50
||||:||||: :|||||
2483 TAGTCATAATCATATCATATCATATCATATCATATCATATCATATTA 2532

50 rgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIlePro... 65
||||:||||: :|||||
2533 ATAACATAATATATCAAAAGAGTCATACCCGTCATTTTCTTTACCAGCT 2582

66 .....AsnArgGlnMetAsnAspGlyLeuGlyGlyAspGlyAspAs 79
||||:||||: :|||||
2583 AATGCTTACCATAGAGAAGTAAACAGCTCTGTAAACCAATAAT..... 2624

79 pMetGluMetPheMetGluGlu..MetArgGluIleArgArgLysLeuArg 95
||||:||||: :|||||
2625 .....TTCTCAACCAATATG.....CACAAATTA 2676

96 GluLeuGlnLeuArgAsnCysLeu 103
||||:||||: |||
2653 AAATTCACCTCTCCGCAACAAATTA 2676

seq_name: /cgn2_6/ptodata/2/1na/6B_COMB.seq:US-09-323-433A-3

seq_documentation_block:
; Sequence 3, Application US/09323433A
; Patent No. 6218512
; GENERAL INFORMATION:
; APPLICANT: Guarente, Leonard P.
; APPLICANT: Austriaco Jr., Nicanor
; APPLICANT: Claus, James J.
; APPLICANT: Cole, Francesca
; APPLICANT: Kennedy, Brian
; TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESCENCE IN
; FILE REFERENCE: 0050.1491-003
; CURRENT APPLICATION NUMBER: US/09/323.433A
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 08/396,001
; PRIOR FILING DATE: 1995-02-28
; PRIOR APPLICATION NUMBER: PCT/US94/09351
; PRIOR FILING DATE: 1994-08-15
; PRIOR APPLICATION NUMBER: US 08/107,408
; PRIOR FILING DATE: 1993-08-16
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 3455
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (663)...(3164)

; OTHER INFORMATION: UTH4
US-09-323-433A-3

alignment_scores:
    Quality: 79.00      Length: 92
    Ratio: 1.362      Gaps: 5
    Percent Similarity: 63.043      Percent Identity: 26.087

alignment_block:
US-09-327-750D-12 x US-09-323-433A-3 ..
Align seg 1/1 to: US-09-323-433A-3 from: 1 to: 3455

17 AsnGlyGluGluAspArgProValGlyGlyGlyGlyGlyGlyHisGlnProAl 33
||||:||||: :|||||
2451 AACACGAAATAAGAACCC.....CATAACAAAAA 2482

33 aGlyAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArgA 50
||||:||||: :|||||
2483 TAGTCATAATCATATCATATCATATCATATCATATCATATCATATTA 2532

50 rgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIlePro... 65
||||:||||: :|||||
2533 ATAACATAATATATCAAAAGAGTCATACCCGTCATTTTCTTTACCAGCT 2582

66 .....AsnArgGlnMetAsnAspGlyLeuGlyGlyAspGlyAspAs 79
||||:||||: :|||||
2583 AATGCTTACCATAGAGAAGTAAACAGCTCTGTAAACCAATAAT..... 2624

79 pMetGluMetPheMetGluGlu..MetArgGluIleArgArgLysLeuArg 95
||||:||||: :|||||
2625 .....TTCTCAACCAATATG.....CACAAATTA 2676

96 GluLeuGlnLeuArgAsnCysLeu 103
||||:||||: |||
2653 AAATTCACCTCTCCGCAACAAATTA 2676

seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-155-888-1

seq_documentation_block:
; Sequence 1, Application US/08155888
; Patent No. 6066623
; GENERAL INFORMATION:
; APPLICANT: Hoffman, Stephen L.
; APPLICANT: Hedstrom, Richard C.
; APPLICANT: Sedegah, Martha
; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE PROTECTIVE
; TITLE OF INVENTION: AGAINST MALARIA, METHODS OF PROTECTION AND VECTOR FOR
; TITLE OF INVENTION: DELIVERING POLYNUCLEOTIDE VACCINES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Naval Medical Res. & Dev. Cmd.
; STREET: Bldg. 1, T-12 8901 Wisconsin Ave.
; CITY: Bethesda
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20889-5606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/155,888
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spevack, A. David
; REGISTRATION NUMBER: 24,743
; REFERENCE/DOCKET NUMBER: N.C. 75.851
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 295-6759
```



```

95 rgGluLeuInLeu.....ArgAsnCysLeu.....Arg 104
||||| :||| :|||
280623 AGCAGCGGCACGTTCGGCCATCAGTCCGGTGCGGCAACCAACGCC 280672
seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-103-840A-2

105 IleLeuMetGlyGluLeuSerAsnHisHls.....AspHls 116
||||| :||| :|||
280673 TTGCCACAGGGGAAGTTCCATGGCATCACCCCGCAGGATCAC 280714
seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-103-840A-1

seq_documentation_block:
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 441529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

alignment_scores:
Quality: 80.00 Length: 164
Ratio: 1.013 Gaps: 8
Percent Similarity: 48.171 Percent Identity: 23.171

alignment_block:
US-09-327-750D-12 x US-09-103-840A-1 ..
Align seg 1/1 to: US-09-103-840A-2 from: 1 to: 4403765

5 HisGlnGluAsnGluMetGluGlnProLeuGlnAsn...GlyGluL 20
||||| :||| :|||
280223 CATCCGCGCGAACAACTTTCGACGAAATCGCTGTAATTCGCGGTGA 280272

20 uAspArgProValGlyGlyGlyGly.....HisGlnProA 33
||||| :||| :|||
280273 CGATGAACCGTTGCGCGCATGAGATTGCGCGTTCCTCCACCCTG 280322

33 laGlyAsnAsnAsnAsnAsnHis.....AsnHisAsnHisAsn 46
||||| :||| :|||
280323 CCGGTGACGGCGGTCTGACCGCGCGTCCAGATCGCGCGCACGATAAC 280372

47 HisHisArgArgGlyGlnAlaArg..... 54
||||| :||| :|||
280373 GATGAACGGGTCTGCTGCGCGGAGCTCGAGCACGTCGGCTTGATCTCGT 280422

55 .....ArgLeuAlaProAsnPheA 61
||||| :||| :|||
280423 TACCGCGCATAGCCCCACCGATTGGCGCGCGCTCGCTCCGGTCAGC 280472

61 rgTrpAlaIleProAsnArgGlnMetAsnAspGlyLeuGlyGly..... 75
||||| :||| :|||
280473 GTGGCGCGCGACCCGGGATCAGCAGGATGCTTCGACGCGTCCCGA 280522

76 .....AspGlyAs 78
||||| :||| :|||
280523 GCTAACAGCAACTCTGGAAGACGCGTCCGGGAAGCGCGCTCGGGCGA 280572

78 pAspMetGluMetPheMetGluGluMetArgGluIleArgArgLysLeuA 95
||||| :||| :|||
280573 TCAGCTCGGCGAGGTACAGCGCGATTCGCGCACGTTTCGACGCGCTGTTG 280622

```

```

STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 80161 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-036-987A-1

```

```

alignment_scores:
  Quality: 81.00      Length: 144
  Ratio: 1.141      Gaps: 5
  Percent Similarity: 49.306      Percent Identity: 23.611
alignment_block:
US-09-327-750D-12 x US-09-036-987A-1/rev ..

```

```

Align seg 1/1 to reverse of: US-09-036-987A-1 from: 1 to: 80161

5 HisGlnGluAsnGluGluMetGluGlnProLeuGlnAsnGlyGluGluAs 21
||||| : : : : : |||
69516 CACCAACTCGGTGATTCCAGAGCATCCACGACGACGCGCGGTCAACA 69467
21 pArgProValGlyGlyGlyGluGlyHisGlnProAlaGlyAsnAsn... 36
: : : : : : : : : : |||||
69466 CCAATCGCGGACACCGCGTTCGGTCACAGGTGGCGGCAACGCGCCC 69417
37 .....AsnAsnAsnHis.....Asn 42
69416 CCCAACAGGCGCGTCCCGCCGGAATCAGCACGGTTCGGTCGGGATCCGC 69367
43 HisAsnHisAsnHis.....HisArgArgGlyG1 52
||||| : : : : : |||
69366 CACAAACCGCGGACGTCGATCTCCCGGCGACCGCAAGCGGCGCG 69317
52 nAlaArgArgLeuAlaPro.....AspPheArgTrpAlaIleP 65
: : : : : : : : : : |||
69316 CTCGCACCGCGCCTTCGCGAGGCCCAACTCGCGTTCATCGACCCCAAC 69267
65 roAsnArgGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGlu 81
: : : : : : : : : : |||
69266 GCACCTCGGCAACGACGCGCATTCGCGGCTTCGCTCGACGTCGCGCAG 69217
82 MetPheMetGluGluMetArgGluIleArgArgLysLeuArgGluGlu 98
: : : : : : : : : : |||
69216 CACAAACTGCGCGGGTTCCTCCACTCGCGGCGACGCAACACCCCAACA 69167
98 nLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHis. 114
| ||| |||
69166 ACGCGCGCATGTGCCAGCTCTCATCCACCGTCACCTCGGAGCGGTGTCACC 69117
115 .....AspHisHisAspGluPheCysLeu 122
||||| : : |||||

```

```

69116 GCTGCACACGTCACACCAACGACGACGTCGCTC 69085
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-370-700-1
seq_documentation_block:
; Sequence 1, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/36987
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 80161
; TYPE: DNA
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-1

```

```

alignment_scores:
  Quality: 81.00      Length: 144
  Ratio: 1.141      Gaps: 5
  Percent Similarity: 49.306      Percent Identity: 23.611
alignment_block:
US-09-327-750D-12 x US-09-370-700-1/rev ..

```

```

Align seg 1/1 to reverse of: US-09-370-700-1 from: 1 to: 80161

5 HisGlnGluAsnGluGluMetGluGlnProLeuGlnAsnGlyGluGluAs 21
||||| : : : : : |||
69516 CACCAACTCGGTGATTCCAGAGCATCCACGACGACGCGCGGTCAACA 69467
21 pArgProValGlyGlyGlyGluGlyHisGlnProAlaGlyAsnAsn... 36
: : : : : : : : : : |||||
69466 CCAATCGCGGACACCGCGTTCGGTCACAGGTGGCGGCAACGCGCCC 69417
37 .....AsnAsnAsnHis.....Asn 42
69416 CCCAACAGGCGCGTCCCGCCGGAATCAGCACGGTTCGGTCGGGATCCGC 69367
43 HisAsnHisAsnHis.....HisArgArgGlyG1 52
||||| : : : : : |||
69366 CACAAACCGCGGACGTCGATCTCCCGGCGACCGCAAGCGGCGCG 69317
52 nAlaArgArgLeuAlaPro.....AspPheArgTrpAlaIleP 65
: : : : : : : : : : |||
69316 CTCGCACCGCGCCTTCGCGAGGCCCAACTCGCGTTCATCGACCCCAAC 69267
65 roAsnArgGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGlu 81
: : : : : : : : : : |||
69266 GCACCTCGGCAACGACGCGCATTCGCGGCTTCGCTCGACGTCGCGCAG 69217
82 MetPheMetGluGluMetArgGluIleArgArgLysLeuArgGluGlu 98
: : : : : : : : : : |||
69216 CACAAACTGCGCGGGTTCCTCCACTCGCGGCGACGCAACACCCCAACA 69167
98 nLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHis. 114
| ||| |||
69166 ACGCGCGCATGTGCCAGCTCTCATCCACCGTCACCTCGGAGCGGTGTCACC 69117
115 .....AspHisHisAspGluPheCysLeu 122
||||| : : |||||

```


1

OM of: US-09-327-750D-12 to: Issued_Patents_NA:* out_format : pfs
Date: Mar 11, 2002 3:36 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framet.p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US09327750/runat_11032002_101154_20340/app_query.fasta.1.1472
-DB=Issued_Patents_NA -Qfmt=fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -GAPOP=10.000 -GAPEXT=0.500
-GAPOP=6.000 -GAPEXT=7.000 -GAPOP=10.000 -GAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blossum62
-TRANS-human40.cdd -LIST=45 -DOALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09327750.@CGN1_1_99 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT_THREADS=1

Search Information block:

Query: US-09-327-750D-12
Query length: 124
Database: Issued_Patents_NA:*
Database sequences: 351203
Database length: 113238999
Search time (sec): 146.090000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US92-06840-1	84.50	144.01	2.27	2353	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-118-319-1	83.00	144.66	2.08	1631	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US92-06840-1	81.50	143.37	2.46	1373	Pate
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-171-969-1	81.00	141.56	3.10	1488	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-036-987A-1	81.00	101.63	519.83	80161	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-370-700-1	81.00	101.63	519.83	80161	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-103-840A-2	80.00	59.49	1.1e-05	4403765	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-103-840A-1	80.00	59.48	1.1e-05	4411529	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-861-464-3	79.00	129.10	15.34	3455	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-396-001-3	79.00	129.10	15.34	3455	
/cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-323-433A-3	79.00	129.10	15.34	3455	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-155-888-1	78.50	123.34	32.11	5552	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-023-173-4	77.00	136.18	6.19	1140	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-387-942C-18	77.00	136.05	6.29	1155	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-07-951-715A-6	77.00	124.59	27.33	3624	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-459-448A-6	77.00	124.59	27.33	3624	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-459-595A-6	77.00	124.59	27.33	3624	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-459-504B-6	77.00	124.59	27.33	3624	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-459-444-6	77.00	124.59	27.33	3624	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-053-549-7	77.00	124.59	27.33	3624	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-317-310A-1	77.00	123.39	31.91	4088	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-945-283-1	77.00	116.13	80.95	8438	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-387-942C-1	77.00	112.12	135.32	12588	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-789-275-1	76.50	127.83	18.05	2373	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-677-862-1	76.50	124.78	26.70	3218	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-252-571-1	76.50	124.78	26.70	3218	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-434-065-1	76.50	124.78	26.70	3218	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-245-295-5	76.50	120.57	45.82	4900	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-481-130-5	76.50	120.57	45.82	4900	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-656-984A-5	76.50	120.57	45.82	4900	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-485-604-5	76.50	120.57	45.82	4900	
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-487-595-5	76.50	120.57	45.82	4900	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-863-790-26	76.50	120.57	45.82	4900	
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-296-749-26	76.50	120.57	45.82	4900	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-481-130-8	76.50	120.21	47.95	5077	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-485-984A-8	76.50	120.21	47.95	5077	
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-485-604-8	76.50	120.21	47.95	5077	

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-487-595-8 - 76.50 120.21 47.95 5077
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-024-020B-8 - 76.50 117.25 70.14 6826
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-356-952-12 - 76.50 98.66 761.02 43676
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-415-751-47 - 76.00 134.65 7.53 1086
/cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-025-691-2 + 76.00 132.90 9.42 1294
seq_documentation_block:
; Sequence 1, Application PC/TUS9206840
; GENERAL INFORMATION:
; APPLICANT: Shi, Yang
; APPLICANT: Seto, Edward
; APPLICANT: Shenk, Thomas
; TITLE OF INVENTION: Y11 TRANSCRIPTION FACTOR AND METHODS OF
; ISOLATING SAME
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
; STREET: 1180 Avenue of the Americas - 7th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-8403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06840
; FILING DATE: 19920814
; CLASSIFICATION:
; CLASSIFICATION: AU 1805
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/746,485
; FILING DATE: 16-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Dennis, Manette
; REGISTRATION NUMBER: 30,623
; REFERENCE/DOCKET NUMBER: M-12594 CIP (1570-8)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 382-0700
; TELEFAX: (212) 382-0888
; TELEX: 236925
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2353 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Hela cells derived from cervical
; TISSUE TYPE: carcinoma
; CELL TYPE: tumor cells
; CELL LINE: Hela
; IMMEDIATE SOURCE:
; LIBRARY: D98/AH-2
; CLONE: p14-1 or pY1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 241..1485
; PCT-US92-06840-1
alignment_scores:
Quality: 84.50 Length: 86
Ratio: 1.798 Gaps: 3
Percent Similarity: 54.651 Percent Identity: 29.070

THIS PAGE BLANK (USPTO)

```
US-09-327-750D-13 x AAC10889  ..
Align seg 1/1 to: AAC10889 from: 1 to: 692

 3 AsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAsnGlyG1 19
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282 AATGTC AACCAAGGAAATGATGAAAAAGATGAAAGGAGCAAGTGWGCTAA 331
 19 uGluAspArgProLeuGly.....GlyGlyGluGlyHisGlnP 32
  ||:|||||:|||||:|||||:|||||:|||||:|||||:
332 TAAAGGGGAGCCCTTGGCCCTACCTTTGRTATGYTRGTGAATACGTGTGTGC 381
 32 roAlaGlyAsnArgArgGly...GlnAlaArgArgLeuAlaProAsnPhe 47
  || |||||:|||||:|||||:|||||:|||||:|||||:
382 CTAGAGGAATCGTAGCGGTTCCGCGCTTAGGCAGCCCATCTCCAGTAT 431
 48 ArgTrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspG1 64
  |||||: ||: |||||: ||: |||||: ||: |||||: ||: |||||
432 AGATGGGATATGATGCATAGCTTTGGAGAACCCACAGGCAAGGATGANAGA 481
 64 yAspAspMetCluIlePheMetGluGluMetArgGluIleArgArgLysL 81
  ||:|||||: ||: |||||: ||: |||||: ||: |||||: ||: |||||
482 AGAATATGGAAAGGATTGGGGAGGAGGTGAGACAGCTGATGGAAAGC 531
 81 euArgGluLeuGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeu 97
  |||||: |||||: |||||: |||||: |||||: |||||: |||||
532 TGAGGGGAAAGCAGTTGAGTCATAGTCTGCGGGCAGTCAGCAGCTGAC... 578
 98 SerAsnHisHisAspHisHisAspGluPheCysLeuMetPro 111
  |||||: |||||: |||||: |||||: |||||: |||||: |||||
579 CCCCTCACCATGACCATCATGATGAGCTTTTGCNNNWTGCC 620
```


Align seg 1/1 to: AAF58581 from: 1 to: 862

```

3 AsnIleHisGlnGluAsn.....GluGluMetGluGlnPrometG1 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
271 AATGTCACACGAGAAATGATGAAAAAGATGAAAGAGGACCAAGTTGCTAA 320
16 nAsnGlyGluGlu...AspArgProLeuGlyGlyGlyHisGlnP 32
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
321 TAAAGGGAGCCCTTGGCCCTACCTTTGAATGTAGTGAATAGTGTGTC 370
32 roLaGlyAsnArgArgGly...GlnAlaArgArgLeuAlaProAsnPhe 47
|||:|||||:|||||:|||||:|||||:|||||:|||||:
371 CTAGAGAAACCGTAGCGGTTCCGGCTTAGGCAGCCCATCTGCAGTAT 420
48 ArgTrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyGlyaspG1 64
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
421 AGATGGGACATATGCTAGCTTGGAGAGCCACAGGCAAGGATGAGAGA 470
64 yAspAspMetGluIlePheMetGluGluMetArgGluIleArgArgLysL 81
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
471 GGAGATATGGAAGGATTTGGGAGGAGGTGACACAGCTGATGAAAGC 520
81 euArgGluLeuGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeu 97
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
521 TGAGGGGAAAAGCAGTTGAGTCATAGTTTGGCGGCAGTCAGCACTGAT... 567
98 SerAsnHisHisAspHisAspGluPheCysLeuMetPro 111
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
568 CCCCTCACCATGACCATCAGATGAGTTTGGCTTATGCCCC 609

```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF59611

seq_documentation_block:

ID AAF59611 standard; cDNA; 898 BP.

XX AAF59611;

XX 24-APR-2001 (first entry)

XX Human cell cycle and proliferation protein CCYPR-22 cDNA, SEQ ID NO:76.

XX Cell cycle and proliferation protein; CCYPR; human; agonist;
 KW antagonist; gene therapy; detection; gene therapy;
 KW transgenic animal disease model; immune disorder;
 KW developmental disorder; cell signalling disorder;
 KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
 KW arteriosclerosis; asthma; allergy; diabetes mellitus;
 KW menstrual cycle disorder; bacterial infection; ss.

XX Homo sapiens.

XX WO200107471-A2.

XX 01-FEB-2001.

XX 21-JUL-2000; 2000WO-US19948.

XX 21-JUL-1999; 99US-0145075.

PR 08-SEP-1999; 99US-0153129.

XX 10-NOV-1999; 99US-0164647.

PA (INCY-) INCYTE GENOMICS INC.

XX Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;

PI Azlmzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;

XX WPI; 2001-112727/12.

XX P-PSDB; AAB60474.

XX Human cell cycle and proliferation proteins and polynucleotides are

PT used to treat, diagnose and prevent immune, developmental and cell

PT signalling disorders and cell proliferative disorders including cancer -

XX

PS Claim 5; Page 181-182; 205pp; English.

XX Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human
 CC cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.
 CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
 CC associated with decreased expression of functional CCYPR, while CCYPR
 CC antagonists are used to treat diseases or conditions associated with
 CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
 CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
 CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
 CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
 CC that specifically bind to CCYPR, and in drug screening methods to
 CC identify compounds that modulate the activity of CCYPR. CCYPR
 CC nucleotides can be used to generate transgenic animal models of human
 CC disease, and can be used in gene therapy in target cells with genetic
 CC abnormalities with respect to the expression of CCYPR for the
 CC treatment or prevention of a disorder associated with CCYPR.
 CC Diseases which can be diagnosed, treated and prevented using CCYPR
 CC proteins, nucleic acids, agonists or antagonists include immune,
 CC developmental and cell signalling disorders, and cell proliferative
 CC disorders including cancer. Specific examples of these disorders
 CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
 CC diabetes mellitus, disorders of the menstrual cycle and infections
 CC caused by bacteria.

XX Sequence 898 BP; 250 A; 186 C; 251 G; 211 T; 0 other;

alignment_scores:

Quality: 175.00 Length: 114
 Ratio: 2.465 Gaps: 4
 Percent Similarity: 62.281 Percent Identity: 41.228

alignment_block:

US-09-327-750D-13 x AAF59611

Align seg 1/1 to: AAF59611 from: 1 to: 898

```

3 AsnIleHisGlnGluAsn.....GluGluMetGluGlnPrometG1 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
299 AATGTCACACGAGAAATGATGAAAAAGATGAAAGAGGACCAAGTTGCTAA 348
16 nAsnGlyGluGlu...AspArgProLeuGlyGlyGlyGlyHisGlnP 32
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
349 TAAAGGGAGCCCTTGGCCCTACCTTTGAATGTAGTGAATAGTGTGTC 398
32 roLaGlyAsnArgArgGly...GlnAlaArgArgLeuAlaProAsnPhe 47
|||:|||||:|||||:|||||:|||||:|||||:|||||:
399 CTAGAGGAAAACCGTAGCGGTTCCGGCTTAGGCAGCCCATCTGCAGTAT 448
48 ArgTrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyGlyaspG1 64
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
449 AGATGGGACATATGCTAGCTTGGAGAGCCACAGGCAAGGATGAGAGA 498
64 yAspAspMetGluIlePheMetGluGluMetArgGluIleArgArgLysL 81
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
499 GGAGATATGGAAGGATTTGGGAGGAGGTGACACAGCTGATGGAAGC 548
81 euArgGluLeuGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeu 97
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
549 TGAGGGGAAAAGCAGTTGAGTCATAGTCTGCGGCGCAGTCAGCACTGAT... 595
98 SerAsnHisHisAspHisAspGluPheCysLeuMetPro 111
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
596 CCCCTCACCATGACCATCAGATGAGTTTGGCTTATGCCCC 637

```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AA03880

seq_documentation_block:

ID AA03880 standard; cDNA; 662 BP.

XX AA03880;

XX

PR 29-JAN-1999; 99US-0118078.
 XX (SEAR) SEARLE & CO G D.
 PA Bunch RT, Curtis SW, Rodi CP, Morris DL;
 XX WPI; 2000-505977/45.
 DR New nucleic acid encoding a carcinogenic biomarker, induced by
 PT phenobarbital treatment of rat hepatocytes, useful for identifying
 PT carcinogenic compounds -
 XX Claim 1; Page 73; 240pp; English.
 XX AAA87080 to AAA87656 represent nucleic acid sequences (N1) encoding a
 CC carcinogenesis biomarkers. The carcinogenesis biomarkers are induced by
 CC treating rat hepatocytes with phenobarbital. The nucleic acids are
 CC useful for identifying carcinogenic compounds. The nucleic acid molecules
 CC can be used to derive probes and/or primers for detecting or inducing
 CC carcinogenesis, respectively.
 XX Sequence 187 BP; 39 A; 48 C; 40 G; 60 T; 0 other;

alignment_scores:
 Quality: 193.00 Length: 35
 Ratio: 5.514 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-327-750D-13 x AAA87147 ..

Align seg 1/1 to: AAA87147 from: 1 to: 187

77 IleArgArgLysLeuArgGluLeuGlnLeuArgAsnCysLeuArgIleLe 93
 |||||
 2 ATCCGAGAAAGCTTAGGGAGCTGCAGTTGAGAAATTCCTGCGTATTCT 51
 93 uMetGlyCluteuSerAsnHisHisAspHisHisAspGluPheCysLeuM 110
 |||||
 52 TATGGGGAGCTCTCTAATCACCGACCATTCAGGATGAATTCGCGCTTA 101
 110 etPro 111
 |||||
 102 TGCCT 106

seq_name: /SDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143

seq_documentation_block:
 ID AAH45143 standard; cDNA; 792 BP.

XX AAH45143;

XX 07-SEP-2001 (first entry)

XX Human brain expressed x-linked protein, hBex, coding sequence.

XX Human; brain expressed x-linked protein; cytostatic; auditory; nontropic;
 KW hBex; dysembryoplasia; hereditary disease; cancer; tumour; deafness;
 KW x-chromosome-binding mental retardation; lissencephalous disease; ss.

XX Homo sapiens.

XX WO200140286-A1.

XX 07-JUN-2001.

XX 27-NOV-2000; 2000WO-CN00502.

XX 30-NOV-1999; 99CN-0124179.

XX (BIOR-) BIOROAD GENE DEV LTD SHANGHAI.

XX

PI Mao Y, Xie Y;
 XX WPI: 2001-397944/42.
 DR P-PSDB; AAB99224.
 XX Isolated human brain-expressed X-linked polypeptide used to diagnose
 PT and treat of dysembryoplasia, hereditary diseases, cancer, tumor,
 PT deafness and x-chromosome-binding mental retardation -
 XX Claim 5; Page 22; 30pp; Chinese.
 XX The present sequence is the coding sequence for a human brain-expressed
 CC x-linked protein (hBex). hBex and its coding sequence are useful in the
 CC diagnosis and treatment of dysembryoplasia, hereditary diseases, cancer,
 CC tumours, deafness, x-chromosome-binding mental retardation and
 CC lissencephalous disease. hBex is also useful for screening mimics,
 CC agonists, or inhibitors, and in peptide fingerprinting identification.
 CC hBex coding sequence can be used as primers or probes, or in producing
 CC gene chips or microarrays.
 XX Sequence 792 BP; 214 A; 172 C; 219 G; 187 T; 0 other;

alignment_scores:
 Quality: 187.50 Length: 113
 Ratio: 2.679 Gaps: 3
 Percent Similarity: 61.947 Percent Identity: 42.478

alignment_block:
 US-09-327-750D-13 x AAH45143 ..

Align seg 1/1 to: AAH45143 from: 1 to: 792

1 MetAlaAsnIleHisGlnGluAsnGluGluMetGluGlnProMetGlnAs 17
 |||||
 212 ATGAAATGCCAACCAAGAAATGAAGAAAGGAGCAAGTTGCTAATAA 261
 17 nGlyGluGlu...AspArgProLeuGlyGlyGlyGlyHisGlnProA 33
 |||||
 262 AGGGAGCCCTTGCCCTCCCTTTGGATGCTGGTGAATACTGTGTGCTA 311
 33 laGlyAsnArgArgGly...GlnAlaArgArgLeuAlaProAsnPheArg 48
 |||||
 312 GAGGAAATCTAGCGGCTTCGCGCTTAGGCAGCCCATCTGCAGTATAGA 361
 49 TrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAs 65
 |||||
 362 TGGGATATGATGATAGCTTAGGCTTGGAGAACACACAGCGAGGATGAGAGA 411
 65 pAspMetGluIlePheMetGluGluMetArgGluIleArgArgLysLeuA 82
 |||||
 412 GAATATGGAAGAGATTGGGGAGGGGTGAGACACCTGATGGAANGCTGA 461
 82 rgGluLeuGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSer 98
 |||||
 462 GGGAAAGACAGTTGAGTCATAGTCTGCGGCGCAGTCAGCACTGAC...CCC 508
 99 AsnHisHisAspHisHisAspGluPheCysLeuMetPro 111
 |||||
 509 CCTCACCATGACCATCATGATGAGCTTTTGGCTTATGCCC 547

seq_name: /SDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH75810

seq_documentation_block:
 ID AAH75810 standard; cDNA; 767 BP.

XX AAH75810;

XX 17-OCT-2001 (first entry)

XX Human x chromosome linked gene expression protein 14 coding sequence.
 DE Human; x chromosome linked gene expression protein 14; cancer;
 XX

286 AGAAGATCCGCCATTGGAGGGGTGAAGCCAGAACCTGGAGAA 335
 35 sn...ArgArgGlyGlnAlaArgLeuAlaProAsnPheArgTrpAla 50
 336 ATATCAGCGGGGGCGAGTTAGGCGACTTGCTCCTAATTTTCGATGGCC 385
 51 IleProAsnArgGlnIleAsnAspGlyMetGlyClyAspGlyAspAspMe 67
 386 ATACCTAATAGGCATATTGAGCAACAATGAAGCG.....AGAGATGATGT 429
 67 tGluIlePheMetGluMetGluMetArgGluIleArgArgLysLeuArgGluL 84
 430 AGAAGGTTTGTAGGCGAGATGATGGAATCAAGAGAAGAACTAGGGAAC 479
 84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
 480 AGCAGATGAGCGACTATATCGCGTTCCTCAACTCTGAACTGACACCAT 529
 101 HisAspHisAspGluPheCysLeuMetPro 111
 530 TATGAC.....TTTGGCTCATACCT 550

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAC85548

seq_documentation_block:

ID AAC85548 standard; cDNA; 1364 BP.

AC AAC85548;

DT 04-JUN-2001 (first entry)

DE cDNA encoding CDIFF-4, Incyte ID No. 1990956CB1.

Human; cell differentiation; CDIFF; agonist; antagonist; epilepsy;
 cell proliferation; Alzheimer's disease; schizophrenia disorder;
 arteriosclerosis; cancer; atherosclerosis; diabetes mellitus; ss.
 Homo sapiens.

Key Location/Qualifiers
 CDS 250..612
 /*tag= a
 /product= "CDIFF-4"

XX WO200119860-A2.

XX 22-MAR-2001.

XX 14-SEP-2000; 2000WO-US25435.

XX 15-SEP-1999; 99US-0154140.

XX 06-DEC-1999; 99US-0169155.

XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Hillman JL, Yue H, Reddy R, Lal P, Shah P, Azimzai Y;
 PI Baughn MR, Lu DAM, Bandman O, Shih LL, Patterson C;

XX WPI: 2001-211447/21.

XX P-PSDB; AAB47126.

XX Isolated polypeptides and polynucleotides involved in cell
 PT differentiation are used for treatment, prevention and diagnosis of
 PT cell proliferative, developmental and neurological disorders e.g.
 PT cancer and Alzheimer's disease -

XX Claim 5; Page 121; 137pp; English.

XX The sequences given in AAC85545-72 encode human polypeptides involved
 CC in cell differentiation (CDIFF). CDIFF polypeptides and agonists of
 CC these are used to treat a disease or condition associated with
 CC decreased expression of functional CDIFF. An antagonist of CDIFF is
 CC used to treat a disease or condition associated with over expression

CC of functional CDIFF. CDIFF polypeptides may be used for the treatment,
 CC prevention and diagnosis of cell proliferative, developmental and
 CC neurological disorders, such as Alzheimer's disease, schizophrenia
 CC disorders, arteriosclerosis, cancer, atherosclerosis, diabetes mellitus
 CC and epilepsy. The CDIFF-4 sequence is homologous to Mus musculus
 CC REX-3. This sequence maps to chromosome 1 within the interval from
 CC 152.2 to 157.4 centiMorgans, to chromosome 3 within the interval from
 CC 157.4 to 158.0 centiMorgans, and to the X chromosome within the interval
 CC from 104.9 to 150.3 centiMorgans.

XX Sequence 1364 BP; 411 A; 269 C; 322 G; 362 T; 0 other;

alignment_scores:

Quality: 277.00 Length: 111

Ratio: 3.298 Gaps: 4

Percent Similarity: 75.676 Percent Identity: 54.955

alignment_block:

US-09-327-750D-13 x AAC85548

Align seg 1/1 to: AAC85548 from: 1 to: 1364

3 AsnIleHisGlnGluAsnGluMetGluGln...PrometGlnAsnG 18

295 AATGCCCAACAAGAAACGAAGGAGGAGGAGGCCGCCACGCAATGA 344

18 yGluLeuAspArgProLeuGlyGlyGlyGlyHisGlnProAlaGlyA 35

345 AGAAGAAATCCGCCATTTGGAGGGGTGAAGGCCAGAACCTGGAGAA 394

35 sn...ArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAla 50

395 ATATCAGCGGGGGCGAGTTAGGCGACTTGCTCCTAATTTTCGATGGCC 444

51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspAspMe 67

445 ATACCTAATAGGCATATTGAGCAACAATGAAGCG.....AGAGATGATGT 488

67 tGluIlePheMetGluMetGluMetArgGluIleArgArgLysLeuArgGluL 84

489 AGAAGGTTTGTAGGCGAGATGATGGAATCAAGAGAAGAACTAGGGAAC 538

84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100

539 AGCAGATGAGCGACTATATCGCGTTCCTCAACTCTGAACTGACACCAT 588

101 HisAspHisAspGluPheCysLeuMetPro 111

589 TATGAC.....TTTGGCTCATACCT 609

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAA87147

seq_documentation_block:

ID AAA87147 standard; DNA; 187 BP.

XX AC AAA87147;

XX 08-JAN-2001 (first entry)

XX Rat hepatocyte carcinogenesis biomarker nucleic acid SEQ ID NO:71.

XX Rat; phenobarbital; carcinogenesis marker; carcinogenesis; detection;
 KW identification; carcinogenic; probe; primer; ds.

XX Rattus norvegicus.

XX WO200044902-A2.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US00503.

|||||
369 ATACCAATAGGCAGATCAATGATGGGATGGTGRAGATGAGATGAT 418
67 tGluilePheMetGluGluMetArgGluileArgArgLysLeuArgGluL 84
|||||
419 GGAATATTTCATGGAGGAGATGAGAGAAATCAGAAGAAATTTAGGGAGC 468
84 euGlnLeuArgAsnCysLeuArgTleLeuMetGlyGluLeuSerAsnHis 100
|||||
469 TGCATTTGAGGAATTTCTCGGTATCTTATGGGGAKC1CTCTAATCAC 518
101 HisAspHisHis 104
|||||
519 CATGACCATCAT 530

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH03517

seq_documentation_block:

ID AAC01004 standard; cDNA; 485 BP.

AC AAC01004;

XX 06-OCT-2000 (first entry)

DT Human secreted protein 5' EST, SEQ ID NO: 1002.

DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping; ss.

XX Homo sapiens.

XX EP1033401-A2.

PN 06-SEP-2000.

PD 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

DR P-PSDB; AAG00998.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 1; SEQ ID 1002; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.

XX Sequence 485 BP; 125 A; 113 C; 159 G; 84 T; 4 other;

alignment_scores:

Quality: 521.00

Length: 99

Gaps: 0

Percent Similarity: 96.970 Percent Identity: 96.970
alignment_block:
US-09-327-750D-13 x AAC01004 ..
Align seg 1/1 to: AAC01004 from: 1 to: 485
6 GlnGluAsnGluGluMetGluGluInProMetGlnAsnGlyGluGluAspAR 22
|||||
187 CAGGAAACGAAGAGATGGAGCAGCCTTGCAGATGGAGAGGAGACCG 236
22 gProLeuGlyGlyGlyGluGlyHisGlnProAlaGlyAsnArgArgGlyG 39
|||||
237 CCCTTTGGAGGAGGTGAAGGCCACAGCCTGCAGAAATCGACGGGAC 286
39 InAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnArgGln 55
|||||
287 AGGCTCGCGACTTGCCTTAATTTTCGATGGGCATACCCAATAGGCAG 336
56 IleAsnAspGlyMetGlyGlyAspGlyAspAspMetGluIlePheMetG 72
|||||
337 ATCAATGATGGGATGGGTGRAGATGGAGATGATATGGAATATTATCGA 386
72 uGluMetArgGluIleArgArgLysLeuArgGluLeuGlnLeuArgAsnC 89
|||||
387 GGAGATCAGAGAAATCAGACAAACTTAGGAGCTGCAKTTGAGGAATT 436
89 ysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisHis 104
|||||
437 GTCTGCGTATCCTTTATGGGGAKCTCTCTAATCACCATGACCATCAT 483

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH03517

seq_documentation_block:

ID AAH03517 standard; cDNA; 865 BP.

XX AAH03517;

XX 26-JUN-2001 (first entry)

XX Human cDNA clone (5'-primer) SEQ ID NO:352.

DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

KW Homo sapiens.

XX EP1074617-A2.

PN 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

XX Claim 1; SEQ ID 352; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesising 5602

PR 07-JUN-1999; 99US-0327750.
XX (UYCO) UNIV COLUMBIA NEW YORK.
XX Sato T;
XX WPI; 2001-061707/07.
XX New p75-neurotrophin receptor-associated cell death executor (NADE) and
XX the gene encoding NADE, useful for modulating the activity of p75NTR
XX and for detecting neurodegenerative diseases -
XX Claim 12; Fig 1; 134pp; English.
XX The present invention relates to a purified polypeptide capable of
XX binding neurotrophin receptor (p75-NTR). The invention is useful for
XX binding and modulating the activity of p75NTR. The peptide mediates
XX NGF-induced apoptosis, which plays an important role in neurogenetic
XX diseases. The peptide of the invention and p75NTR are useful for
XX inhibiting NF-kappaB activation in a cell or a subject, for inducing
XX caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase
XX and fragment nuclear DNA in a cell by co-expression of (1) and p75-NTR.
XX Sequence 700 BP; 177 A; 188 C; 203 G; 132 T; 0 other;

alignment_scores:
Quality: 564.50 Length: 124
Ratio: 5.086 Gaps: 1
Percent Similarity: 89.516 Percent Identity: 82.258

alignment_block:

US-09-327-750D-13 x AAF23528 ..

Align seg 1/1 to: AAF23528 from: 1 to: 700

1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
177 ATGGCCAATCTCCACCAAGGAAGAGAGCTGGAGCAGCCCTGCAGAA 226
17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyGlyHisGlnProAlaG 34
227 TGGACAGGAAGACCGCCCTGTGGAGGAGGTGAGGGCCACGAGCTGCTG 276
34 Lysn.....AtgArg 37
277 CAACACACACACACACACACACACATACACACACACACACCGAAGA 326
38 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 54
327 GGCAGGCTCGCGACTTGCCCTAACTCCGATGGGCCATTCCCAACAG 376
54 gGlnIleAsnAspGlyMetGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 71
377 GCAGATCAATCAGGGGTGGGTGGAGATGAGATGATGATGATGATGATGAT 426
71 etGluGluMetArgGluIleArgArgLysLeuArgGluLeuGlnLeuArg 87
427 TGGAGGAGATGAGAGAGATCCGAGAAAGCTTAGGGAGCTACAGCTGAGA 476
88 AsnCysLeuArgIleLeuMetGlyGlyLeuSerAsnHisHisAspHisH 104
477 AATTGTCTACGCATCCTTATGGGGAGCTGCTTAACCAACCAACCAACCA 526
104 saspGluPheCysLeuMetPro 111
527 TGATGAATTCGTGCTTATGCT 548

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT: AAC01005

seq_documentation_block:

ID AAC01005 standard; cDNA; 532 BP.

XX

AC AAC01005;
XX 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 1003.
DE Human secreted protein 5' EST, SEQ ID NO: 1003.
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX EP1033401-A2.
PN 06-SEP-2000.
PD 21-FEB-2000; 2000EP-0200610.
XX 26-FEB-1999; 99US-0122487.
PR (GEST) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI WPI; 2000-500381/45.
XX P-PSDB; AAG00999.
DR New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
PS Claim 1; SEQ ID 1003; 71pp + CD-ROM; English.
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified within the
XX sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX derived from 30 different tissues. EST sequences usually correspond
XX mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX well suited for isolating cDNA sequences derived from the 5' ends of
XX mRNAs and even in those cases where longer cDNA sequences have been
XX obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
XX mRNAs with intact 5' ends and can therefore be used to obtain full length
XX cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
XX gene therapy and chromosome mapping procedures. They are used to obtain
XX upstream regulatory sequences and to design expression and secretion
XX vectors.
XX Sequence 532 BP; 151 A; 121 C; 163 G; 91 T; 6 other;

alignment_scores:

Quality: 537.00 Length: 104
Ratio: 5.424 Gaps: 0
Percent Similarity: 95.192 Percent Identity: 95.192

alignment_block:

US-09-327-750D-13 x AAC01005 ..

Align seg 1/1 to: AAC01005 from: 1 to: 532

1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
219 ATGGCAATATTCACCAAGGAAGAGAGATGGAGCAGCCCTATGCAGAA 268
17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyGlyHisGlnProAlaG 34
269 TGGAGAGGAAGACCGCCCTTTGGAGGAGGTGAAGGCCACGAGCTGCAG 318
34 LysAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAla 50
319 GAAATGACGGGAMAGGCTCGCCGAYTTGCCCTTAATTTTCGATGGGCC 368
51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspPme 67

|||||
412 CAAATCGACGGGACAGCTCGCGACTGCCCCCTAAATTTTCGATGGGCC 461
51 IleProAsnArgGlnIleAsnAspGlyMetClyGlyAspGlyAspPme 67
|||||
462 ATACCCCAATAGGCAGATCAATGATGGATGGGTGGAGATGGAGATGAT 511
67 tGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGluL 84
512 GGAATATTTCATGCGAGGAGATGAGAGAAATCAGAAGAACTTAGGAGC 561
84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
|||||
562 TGCAGTTGAGGAATGTCTGGTATCTTATGGGGAGCTCTCTAATCAC 611
101 HisAspHisHisAspGluPheCysLeuMetPro 111
612 CATGACCATCATGATGAATTTTGCCCTTATGCCCT 644

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:AAF21748

seq_documentation_block:

ID AAF21748 standard; DNA; 917 BP.

AC AAF21748;

XX 27-MAR-2001 (first entry)

XX Human breast and ovarian cancer associated antigen gene SEQ ID 135.
KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neutropic; neurprotection; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antinflammatory; antiulcer; vulnerary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease; ds.
XX

OS Homo sapiens.

XX WO200055173-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05881.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-611515/58.

XX P-PSDB; AAB58845.

PT New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases -
XX

XX Claim 1; Page 581-582; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
XX proteins AAB58711 - AAB59128. The DNA and protein sequences are
XX associated with breast and ovarian cancer. Included in the invention are
XX sequences AAF22032 - AAF22040 and AAB59129 which are used in the
XX isolation and characterisation of the DNA and protein sequences of the
XX invention. The breast and ovarian cancer associated DNA, protein, agonist
XX or antagonist sequences exhibit cytostatic; immunosuppressive;
XX neutropic; neuroprotective; antiviral; antiallergic; hepatotropic;
XX antidiabetic; antinflammatory; antiulcer; vulnerary; anticonvulsant;
XX antibacterial; antifungal; antiparasitic and cardiant activity. The

CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and agonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemias; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.
XX

SQ Sequence 917 BP; 228 A; 203 C; 270 G; 211 T; 5 other;

alignment_scores:

Quality: 582.00 Length: 106
Ratio: 5.491 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-13 x AAF21748 ..

Align seg 1/1 to: AAF21748 from: 1 to: 917

6 GlnGluAsnGluGluMetGluGlnProMetGlnAsnGlyGluGluAsp 22
|||||
306 CAGCAAAACGAAGAGATGGAGCGCTATGCAGAAATGGAGAGAACCG 355
22 gProLeuGlyGlyGlyGluGlyHisGlnProAlaGlyAsnArgArgGlyG 39
|||||
356 CCCCTTTGGGAGGAGGTGAAGGCCACCGCTGCAGAAATCGACGGGAC 405
39 InAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnArgGln 55
|||||
406 AGGCTCCCGACCTTGCCCTTAATTTTCATGGGCCATACCCCAATAGGCG 455
56 IleAsnAspGlyMetGlyGlyAspGlyAspMetGluIlePheMetG 72
|||||
456 ATCAATGATGGGATGGGTGGAGATGGAGATGATGGAATATTTCATGGA 505
72 uGluMetArgGluIleArgArgLysLeuArgGluLeuGlnLeuArgAsnC 89
|||||
506 GGAGATGAGAGAAATCAGAAGAAACTTAGGGAGCTGCAGTTGAGGAATT 555
89 ysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisHis 105
|||||
556 GTCTGCGTATCTTATGGGGAGGCTCTTAATCACCATGACCATCATGAT 605
106 GluPheCysLeuMetPro 111
|||||
606 GAATTTTGCCCTTATGCCCT 623

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF23528

seq_documentation_block:

ID AAF23528 standard; DNA; 700 BP.

XX AC AAF23528;

XX DT 22-MAR-2001 (first entry)

XX Mouse NADE DNA.

XX Neurotrophin receptor; p75-NTR; NCF-induced apoptosis;
KW neurogenetic disease; NF-kappaB; ds.

XX Mus sp.

XX WO200075278-A2.

XX 14-DEC-2000.

XX 07-JUN-2000; 2000WO-US15621.

XX

OM of: US-09-327-750D-13 to: N_Geneseq_1101:* out_format : pfs
Date: Mar 11, 2002 3:42 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:
-MODEL=frame+ p2n_model -DEV=xlh
-Q=Cgn2_1/USPTO_spool/US09327750/runat_11032002_101154_20362/app_query.fasta_1.1472
-DB=N_Geneseq_1101 -Qfmt=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-GAPOP=6.000 -GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LISN=45 -DOCLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09327750_CGn1_1_330 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-13
Query length: 111
Database: N_Geneseq_1101:*
Database sequences: 930621
Database length: 428662619
Search time (sec): 355.560000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF23529 +		609.00	1187.76	891	4.4e-58
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAF21748 +		582.00	1134.20	917	4.2e-55
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF23528 +		564.50	1102.28	700	2.5e-53
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAF23528 +		537.00	1050.66	532	1.9e-50
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AA01005 +		521.00	1019.98	485	9.8e-49
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AA01004 +		527.00	532.88	865	1.3e-21
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AA03517 +		277.00	529.49	1229	2.0e-21
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH13750 +		277.00	528.48	1364	2.3e-21
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AA085548 +		193.00	381.92	187	3.4e-13
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH47147 +		175.50	357.12	792	8.2e-12
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143 +		187.00	332.76	767	1.9e-10
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH58581 +		175.00	331.63	862	2.1e-10
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH59611 +		175.00	331.23	898	2.3e-10
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AA03880 +		169.00	322.34	662	7.1e-10
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AA03889 +		169.00	321.91	692	7.4e-10
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AA03879 +		169.00	321.83	698	7.5e-10
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH82586 +		168.00	324.63	426	5.3e-10
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH60367 +		162.00	306.02	858	5.7e-09
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AA01003 +		108.00	206.81	401	0.0019
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH72409 +		105.00	190.32	1197	0.0159
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH14082 +		102.50	196.05	397	0.0076
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH135463 +		102.50	196.05	397	0.0076
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH103935 +		102.50	196.05	397	0.0076
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH121436 +		91.50	177.55	285	0.0819
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH16725 +		91.50	177.55	285	0.0819
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH107131 +		91.50	177.55	285	0.0819
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH68557 +		91.50	163.90	1170	0.4716
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH157960 +		91.50	163.26	1250	0.5119
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AA098961 +		91.50	162.81	1309	0.5421
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH159746 +		91.50	161.17	1551	0.6889
/SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ63293 +		88.50	135.01	12588	19.17
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH81469 +		87.50	117.80	60873	174.20
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH21610 +		87.50	100.90	349980	1.5e+03
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH81490 +		87.50	87.24	1437668	8.8e+03
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAH19956 +		84.50	151.53	1007	2.30
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH34631 +		82.50	128.35	7365	45.02
/SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT:AAH37051 +		82.00	132.67	4255	25.88
/SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT:AAH44922 +		80.00	146.02	711	4.67
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAH87356 +		80.00	140.62	1243	9.34
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAH64101 +		80.00	140.62	1243	9.34

/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AA04966 + 80.00 140.62 9.34 1243
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AA087357 + 80.00 140.40 9.61 127
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AA064099 + 80.00 140.40 9.61 127
/SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AA004964 + 80.00 140.40 9.61 127
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:AAH11328 + 79.00 149.09 3.15 422

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAF23529

seq_documentation_block:

ID AAF23529 standard; DNA; 891 BP.

XX

AC AAF23529;

DT 22-MAR-2001 (first entry)

XX

DE Human NADE DNA.

KW Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;

KW neurogenetic disease; NF-kappaB; ds.

XX

OS Homo sapiens.

XX

PN WO200075278-A2.

PD 14-DEC-2000.

PF 07-JUN-2000; 2000WO-US15621.

XX

PR 07-JUN-1999; 9905-0327750.

XX

PA (UYCO) UNIV COLUMBIA NEW YORK.

XX

PI Sato T;

XX

DR WPI; 2001-061707/07.

XX

PT New p75-neurotrophin receptor-associated cell death executor (NADE) and

PT the gene encoding NADE, useful for modulating the activity of p75NTR

PT and for detecting neurodegenerative diseases -

XX

PS Disclosure: Fig 1; 134pp; English.

XX

CC The present invention relates to a purified polypeptide capable of

CC binding neurotrophin receptor (p75-NTR). The invention is useful for

CC binding and modulating the activity of p75NTR. The peptide mediates

CC NGF-induced apoptosis, which plays an important role in neurogenetic

CC diseases. The peptide of the invention and p75NTR are useful for

CC inhibiting NF-kappaB activation in a cell or a subject, for inducing

CC caspase-2 and caspase-3 activity to cleave poly (Adp-ribose) polymerase

CC and fragment nuclear DNA in a cell by co-expression of (1) and p75-NTR.

XX

SQ Sequence 891 BP; 251 A; 182 C; 224 G; 234 T; 0 other;

alignment_scores:

Quality: 609.00 Length: 111

Ratio: 5.486 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-13 x AAF23529

Align seg 1/1 to: AAF23529 from: 1 to: 891

1 MetAlaAsnIleHisGlnGluAsnGluGluMetGluGlnProMetGlnAs 17

|||||

312 ATGCCAAATATTCCACGAGAAACGAAGATGAGCAGCCTATGCAGAA 361

17 nGlyGluGluAspArgProLeuGlyGlyGlyGluGlyHisGlnProAla 34

|||||

362 TGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 411

34 LyAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAla 50

THIS PAGE BLANK (USPTO)


```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/459,595
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; OTHER INFORMATION: optimized cryiB"
; OTHER INFORMATION: /note= "Disclosed in Figure 6."
;
US-08-459-504B-6

alignment_scores:
  Quality: 75.50      Length: 147
  Ratio: 1.198      Gaps: 7
  Percent Similarity: 42.857      Percent Identity: 25.850

alignment_block:
US-09-327-750D-13 x US-08-459-504B-6
...
Align seg 1/1 to: US-08-459-504B-6 from: 1 to: 3624

5 HisGlnGluAsnGluMetGluGlnPrometGlnAsnGlyGluAs 21
|||||: |||:|||||: |||:|||||: |||:|||||:
2121 CACCACGACGACGACGACGACGACGACGACGACGACGAC 2167
|||||: |||:|||||: |||:|||||: |||:|||||:
21 parGProLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 38
|||||: |||:|||||: |||:|||||: |||:|||||:
2168 ACGGCTGTGGTGGGCGAGGAGGAGGAGGAGGAGGAGGAGG 2217
|||||: |||:|||||: |||:|||||: |||:|||||:
38 lyGlnAla.....ArgArgLeuAlaPro 45
|||||: |||:|||||: |||:|||||: |||:|||||:
2218 TTCAGGAGGAACTACGTGACCTGCCCGGACGCTTCAACGAG 2267
|||||: |||:|||||: |||:|||||: |||:|||||:
45 AsnPhe.....ArgtrpAl 50
|||||: |||:|||||: |||:|||||: |||:|||||:
2268 CACCTACCTGTACCAGAACGCGGAGGAGGAGGAGGAGGAGG 2317
|||||: |||:|||||: |||:|||||: |||:|||||:
50 atleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGly 65
|||||: |||:|||||: |||:|||||: |||:|||||:
2318 GCTACCAAGTGGCGGCTACATCGAGGAGGAGGAGGAGGAGG 2367
|||||: |||:|||||: |||:|||||: |||:|||||:
66 .....AspMet 67
|||||: |||:|||||: |||:|||||: |||:|||||:
2368 CTGATCCGCTACACCCCAAGCAGCAGACCTGGAGCTGCCCGG 2417
|||||: |||:|||||: |||:|||||: |||:|||||:
68 GluIlePheMetGluMetGluMetArgGluIleArgArgGlyLe 84
|||||: |||:|||||: |||:|||||: |||:|||||:
2418 GAGCCTGTGGCCTGAGGCTGAGGAGGAGGAGGAGGAGGAGG 2467
|||||: |||:|||||: |||:|||||: |||:|||||:
84 uGln....LeuArgAsnCysLeuArgIle..... 92
|||||: |||:|||||: |||:|||||: |||:|||||:

```

```

2468 CCAACCGCTGCGCCCGCCACTTCAGTGGAAACCCGACCTGACG 2517
93 ..LeuMet.GlyGluLeuSerAsnHisHisAspHisHis 104
|||||: |||:|||||: |||:|||||: |||:|||||:
2518 TGCCGCGACGCGAGAGTGGCGCCACCCACGACGACCCAC 2556
seq_name: /cgn2_6/ptodata/2/1na/6A_COMB.seq:US-08-459-444-6
seq_documentation_block:
; Sequence 6, Application US/08459444A
; Patent No. 6121014
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; Desai, Nalini M.
; Lewis, Kelly S.
; Kramer, Vance C.
; Warren, Gregory W.
; Evola, Stephen V.
; Crossland, Lyle D.
; Wright, Martha S.
; Merlin, Ellis J.
; Launis, Karen L.
; TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED
; NUCLEIC ACID CODING SEQUENCE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 6121014artis Agribusiness Biotechnology Research, Inc.
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,444A
; FILING DATE: 02-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-18805/P1/CGC1577/CIP/DIV6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; optimized cryiB"
; /note= "Disclosed in Figure 6."
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-459-444-6

alignment_scores:

```

21 pArgProLeuGlyGlyGlyGluGluGlyHisGlnProAlaGlyAsnArgArgG 38
 2168 ACGCGTGTGGGGCAGCAGAGAACATCACCATCCAGGAGGCGCAGCAGGTG 2217
 38 LysGlnAla.....ArgArgLeuAlaPro 45
 2218 TTCAAGGAGAACTACGTCACCGTCGCCGGCACCCTTCAACGAGTGCTACCC 2267
 46 AsnPhe.....ArgTyrPAl 50
 2268 CACCTACCTGTACCAAGAAGATCGCGAGAGCGAGCTGAAGGCCCTACACCC 2317
 50 alleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspClyAsp.... 65
 2318 GCTACCAGCTCGCGGCTACATCTCAGGAGACGCCAGGACCTGGAGCTTAC 2367
 66AspMet 67
 2368 CTGATCCGCTACACGCCACGACGAGACCTGGACGTCGCCGGCACCGA 2417
 68 GluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGluLe 84
 2418 GAGCCTGTGGCCCTTGAGCGTGGAGAGCCCATCGCGCGCTGCGCGCAGC 2467
 84 uGln...LeuArgAsnCysLeuArgIle..... 92
 2468 CCACCGCTGGCGCCCGCCACCTTCGAGTGGGAACCCGACCTGGAGCTGCAGC 2517
 93 ..LeuMet.GlyGluLeuSerAsnHisHisAspHisHis 104
 2518 TGCCCGCAGCGAGCAAGCTGCGGCCACACACAGCCACAC 2556

```

seq_documentation_block:
  ; Sequence 6, Application US/08459504B
  ; Patent No. 6075185
  ; GENERAL INFORMATION:
  ; APPLICANT: Koziel, Michael G.
  ; APPLICANT: Desai, Nalini M.
  ; APPLICANT: Lewis, Kelly S.
  ; APPLICANT: Kramer, Vance C.
  ; APPLICANT: Warren, Gregory W.
  ; APPLICANT: Evola, Stephen V.
  ; APPLICANT: Crossland, Lyle D.
  ; APPLICANT: Wright, Martha S.
  ; APPLICANT: Merlin, Ellis J.
  ; APPLICANT: Launis, Karen L.
  ; APPLICANT: Rothstein, Steven J.
  ; APPLICANT: Bowman, Cindy G.
  ; APPLICANT: Dawson, John L.
  ; APPLICANT: Dunder, Erik M.
  ; APPLICANT: Pace, Gary M.
  ; APPLICANT: Suttie, Janet L.
  ; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
  ; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
  ; NUMBER OF SEQUENCES: 94
  ; CORRESPONDENCE ADDRESS:
  ; ADDRESSEE: No. 6075185artis Corporation
  ; STREET: 3054 Cornwallis Road
  ; CITY: Research Triangle Park
  ; STATE: NC
  ; COUNTRY: USA
  ; ZIP: 27709
  ; COMPUTER READABLE FORM:
  ; MEDIUM TYPE: Floppy disk
  ; COMPUTER: IBM PC compatible
  ; OPERATING SYSTEM: PC-DOS/MS-DOS
  ; SOFTWARE: PatentIn Release #1.0, Version #1.30
  ; CURRENT APPLICATION DATA:
  ; APPLICATION NUMBER: US/08/459,504B
  ; FILING DATE:
  ; CLASSIFICATION:

```

```

alignment_scores:
  Quality: 75.50      Length: 147
  Ratio: 1.198      Gaps: 7
  Percent Similarity: 42.857  Percent Identity: 25.850

alignment_block:
  US-09-327-750D-13 x US-08-459-595A-6  ..

Align seg 1/1 to: US-08-459-595A-6  from: 1 to: 3624

      5  HisGlnGluAsnGluClnuMetClnGlnPrometGlnAsnGlyClnuAs  21
      |||||:  |||:|||||:|||||:
2121  CACC AACGAGGACGACCACTTCACGAG...CATCCACGACGAGCAGGC  2167

```



```

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Other
US-09-179-558-62

alignment_scores:
    Quality: 75.50      Length: 103
    Ratio: 1.452      Gaps: 5
    Percent Similarity: 50.485      Percent Identity: 25.243

alignment_block:
US-09-327-7500D-13 x US-09-179-558-62/rev ..

Align seg 1/1 to reverse of: US-09-179-558-62 from: 1 to: 936

3 AsnIleHISgInGluAsnGlucluMetGluGln.....Pr 14
: : : : : : : : : : : : : : : : : : : : : : : : :
730 AGCCATGTCATCATGACCCCGGGTTCCAAACAGCTCCCGCAACACCA 681

14 oMetGlnAsnGlyGluGluAspArgProLeuGlyGlyGluGlyHisG 31
: : : : : : : : : : : : : : : : : : : : : : : : :
680 GGTACGCGAGGAGAGATGAGGCCCATGTGGCGCGGTGATGGACAGC 631

31 lnProAlaGlyAsnArgArgGlyGlnAlaArgLeuAlaProAsnPh 47
: : : : : : : : : : : : : : : : : : : : : : : : :
630 ACGGCCACGGCAGGCCCGGGGACAACCTGCTGGCA..... 590

48 ArgTrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyGly 64
: : : : : : : : : : : : : : : : : : : : : : : : :
589 .....GAGTCTCTCGCGTCGGCGGAGGAGGTGGAGC 558

64 yAspAspMetGluIlePheMetGluMetGluMetArgGluIleArg 81
: : : : : : : : : : : : : : : : : : : : : : : : :
557 CATCATGATCGGATGTG.....CTGCACGCGCTTCTCTCAAGC 520

81 euArgGluLeuGlnLeuArgAsnCysLeuArgIleLeuMetClyclu 97
: : : : : : : : : : : : : : : : : : : : : : : : :
519 TC.....CTTAGTAAATTCATCCAG..... 500

98 SerAsnHis 100
: : : : :
499 ...AACCAC 494

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-951-715A-6

seq_documentation_block:
; Sequence 6, Application US/07951715A
; Patent No. 5625136
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne

```


2069 CGAGCCCTCCGCGCGCGCGCGCGCTTC.....CACCAACCGCG 2112
 49 pAlaIleProAsnArgGlnIleAsn.....AspGlyMetGlyC 62
 2113 CCACCTCCGCGCGCGCGCGCGCACCTCCGCGCGCGGATCCGCGCGGAG 2162
 62 lyAspGlyAspAspMetGlu 68
 2163 GAGAAGGAGAGGAGGAGGAG 2182

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-179-558-64

```
seq_documentation_block:
; Sequence 64, Application US/09179558
; Patent No 6180612
; GENERAL INFORMATION:
; APPLICANT: Hockensmith, Joel W.
; APPLICANT: Muthuswami, Rohini
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TARGETING DNA METABOLIC PROCESSES USING
; AMINOGLYCOSIDE DERIVATIVES
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESS: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/179,558
; FILING DATE: 27-OCT-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 09/060,470
; FILING DATE: 15-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 60/063,898
; FILING DATE: 31-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 9426-005-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)7909090
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 892 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
US-09-179-558-64
```

alignment_scores:
 Quality: 75.50 Length: 103
 Ratio: 1.452 Gaps: 5
 Percent similarity: 50.485 Percent Identity: 25.243

alignment_block:

US-09-327-750d-13 x US-09-179-558-64/rev ..

Align seg 1/1 to reverse of: US-09-179-558-64 from: 1 to: 892

3 AsnIleHisGlnGluAsnGluGluMetGluGln.....Pr 14

```
686 AGCCTGCATCAGCACCCCGGTTCCAAACACGCTCCCCGAACACCA 637
14 oMetGlnAsnGlyGluGluAspArgProLeuGlyGlyGlyGlyHisG 31
636 GGTACGCCGAGGAGAAGATGAGGCCCATGTTGGCGCGGTGATGACAGC 587
31 lnProAlaGlyAsnArgArgGlyGlnAlaArgArgLeuAlaProAsn 47
586 ACGGCCACGCGCGCGCGCGGCAACTGAACTGCTGTGCA..... 546
48 ArgTrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAsp 64
545 .....GAGGTCTCTCGCGGTGCGCGGAGGAGGTGAGC 514
64 yAspAspMetGluIlePheMetGluGluMetArgGluIleArgArgLys 81
513 CATCGATACGATGTG.....CTGCACGCGCTTCTCTCAAGC 476
81 euArgGluLeuGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeu 97
475 TC.....CTTAGTAATTGCATCCAG..... 456
98 SerAsnHis 100
455 ....AACCAC 450
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-179-558-62
seq_documentation_block:
; Sequence 62, Application US/09179558
; Patent No 6180612
; GENERAL INFORMATION:
; APPLICANT: Hockensmith, Joel W.
; APPLICANT: Muthuswami, Rohini
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TARGETING DNA METABOLIC PROCESSES USING
; AMINOGLYCOSIDE DERIVATIVES
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESS: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/179,558
; FILING DATE: 27-OCT-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 09/060,470
; FILING DATE: 15-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 60/063,898
; FILING DATE: 31-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 9426-005-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)7909090
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 892 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other
US-09-179-558-64
```


.....

77 leArgArgLysLeuArgGluLeuG

928 ACCAGTGCATCCCCCAATACTCCTTCACCCGGGCTCGACAGCGTTTCTGAG 977
77 leArgArgLysLeuArgGluLeuGlnLeuArgAsnCysLeuArgIleLeu 93

```

3137 TGGANGCGGGGAATGAGCACGATCCGGAGGAGATGAAGGAGGCGAGCTAC 3188
41  ArgArgLeuAlaIleProAsnPheArgTrpAlaIle..... 51
   ::::: |||::: |||:::
3187 CATGCCACACCCGCCCAAGTACAGACGGGCTGGTTTATGCTCCTCCGCA 3236
52  .ProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspMetG 68
   ::::: |||::: |||::: |||::: |||
3237 TCCGGATCAAGAGCGCGCCTCCGGATCGGGATCGGAT..... 3273
68  luIlePheMetGluMetArgGluArgGluIleAArgArgLysLeuArgGluLeu 84
   |||::: |||::: ::::: |||::: |||::: |||::: |||
3274 ..ATCATGTGGATGGCGGCTACAATTGGCAGGTACAAGTCAAGGAGCTC 3321
85  GlnLeuArgAsnCysLeuArgIle.....LeuMetGlyGly 96
   ::::: |||::: |||::: |||::: |||::: |||::: |||
3322 AACATGCAGCGCAACATACGATGTCAGTCCCGACGACACATGATGTCCCA 3371
96  uLeuSerAsnHisHisAspHisHis 104
   ::::: |||::: |||::: |||::: |||::: |||
3372 CTATTCCGCGCATCATCCGCCCAT 3396

```

```
Percent Similarity: 55.696      Percent Identity: 36.709

alignment_block:
US-09-327-750D-13 x US-08-387-942C-1      ..
Align seg 1/1 to: US-08-387-942C-1 from: 1 to: 12588

18 GlyGluGluAspArgProLeuGlyGlyGlyGluGlyHisGlnProAlaGl 34
|||||::: ||||| ||||| ::: |||||
2298 GCGCATGCAGCGCCGCCACGGCGGGCGCGCACCGCTACTGTCCGG 2347
34 yAsnArg...ArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpA 50
|:::|| | ||||| ||||| ||| |||||
2348 CCGGCGAATATCGGTTCAGCGCGCGGAGAGCC.....TTCGATGGT 2391
50 laile.....ProAsnArgGlnIleAsnAspGlyMet 60
||| |||::: ::|||
2392 TGTCTGCACCATCAAGACAACGTCCTCATATCGTCGGCGCGGATGGCGA 2441
61 GlyGlyAspGlyAspMetGluIlePheMetGluMetArgGlu.. 76
||||| ||| :::: ||::: ||:::
2442 GACGGTGATCAAGATGCTGACGGCTGGACGCAGCAACGTACC GG CATG 2491
77 .....IleArgArgLysLeuArgGluLeuGln 85
::::: ||||| :::::||:::
2492 TGC GCTCGCGCCTACGCGCAGGAACAGCAACTCGG 2528

seq_name: /cgn2_6/ptodata/2/ina/PCUTS_COMB.seq; seq:PCT-US96-02331-12
seq_documentation_block:
; Sequence 12, Application PC/TUS9602331
; GENERAL INFORMATION:
; APPLICANT: The Board of Trustees of the Leland Stanford Ju
; APPLICANT: University
; APPLICANT: Board of Regents, The University of Texas Syst
; TITLE OF INVENTION: Methods and Compositions for Altering
; TITLE OF INVENTION: Sexual Behavior
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/02331
; FILING DATE: 09-FEB-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,495
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 8600-0153.41
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1244 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

```

; TOPOLOGY: UNKNOWN
; MOLECULE TYPE: CDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO

```


553 TATATGGCTTCCAACTCTGAACTGACACCACTATTATGAC..... 594
118 pGluPheCysLeuMetPro 124
|||||.....
595TTTGGCTCATACCT 609

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:AAA87147

seq_documentation_block:

ID AAA87147 standard; DNA; 187 BP.

XX AC AAA87147;

XX DT 08-JAN-2001 (first entry)

XX DE Rat hepatocyte carcinogenesis biomarker nucleic acid SEQ ID NO:71.

XX KW Rat; phenobarbital; carcinogenesis marker; carcinogenesis; detection;

XX KW Identification; carcinogenic; probe; primer; ds.

XX OS Rattus norvegicus.

XX PN WO200044902-A2.

XX PD 03-AUG-2000.

XX PF 28-JAN-2000; 2000WO-US00503.

XX PR 29-JAN-1999; 99US-0118078.

XX PA (SEAR) SEARLE & CO G D.

XX PI Bunch RT, Curtis SW, Rodi CP, Morris DL;

XX DR WPI; 2000-505977/45.

XX PT New nucleic acid encoding a carcinogenic biomarker, induced by
phenobarbital treatment of rat hepatocytes, useful for identifying

XX PT carcinogenic compounds -

XX PS Claim 1; Page 73; 240pp; English.

XX CC AAA87080 to AAA87656 represent nucleic acid sequences (N1) encoding a
carcinogenesis biomarkers. The carcinogenesis biomarkers are induced by
treating rat hepatocytes with phenobarbital. The nucleic acids are
useful for identifying carcinogenic compounds. The nucleic acid molecules
can be used to derive probes and/or primers for detecting or inducing
carcinogenesis, respectively.

XX SQ Sequence 187 BP; 39 A; 48 C; 40 G; 60 T; 0 other;

alignment_scores:
Quality: 193.00 Length: 35
Ratio: 5.514 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-12 x AAA87147

Align seg 1/1 to: AAA87147 from: 1 to: 187

90 IleArgArgLysLeuArgGluLeuGlnLeuArgAsnCysLeuArgIleLe 106

|||||.....
2 ATCCGGAGAAGCTTAGGGAGCTGCAGTTGAGAAATGTCGTATTCT 51

106 uMetGlyGluLeuSerAsnHisHisAspHisHisAspGluPheCysLeu 123

|||||.....
52 TATGGGGAGCTCTCTAATCACCAGCATCAGATGAATTCGTGCTTA 101

123 etPro 124

|||||

102 TGCCT 106

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAH45143

seq_documentation_block:

ID AAH45143 standard; cDNA; 792 BP.

XX AC AAH45143;

XX DT 07-SEP-2001 (first entry)

XX DE Human brain expressed X-linked protein, hBex, coding sequence.

XX KW Human; brain expressed X-linked protein; cytostatic; auditory; nontropic;

XX KW hBex; dysembryoplasia; hereditary disease; cancer; tumour; deafness;

XX KW X-chromosome-binding mental retardation; lissencephalous disease; ss.

XX OS Homo sapiens.

XX PN WO200140286-A1.

XX PD 07-JUN-2001.

XX PF 27-NOV-2000; 2000WO-CN00502.

XX PR 30-NOV-1999; 99CN-0124179.

XX PA (BIOR-) BIOROAD GENE DEV LTD SHANGHAI.

XX PI Mao Y, Xie Y;

XX DR WPI; 2001-397944/42.

XX P-PSDB; AAB99224.

XX PT Isolated human brain-expressed X-linked polypeptide used to diagnose
and treat of dysembryoplasia, hereditary diseases, cancer, tumor,

XX PT deafness and X-chromosome-binding mental retardation -

XX PS Claim 5; Page 22; 30pp; Chinese.

XX CC The present sequence is the coding sequence for a human brain-expressed
X-linked protein (hBex). hBex and its coding sequence are useful in the
diagnosis and treatment of dysembryoplasia, hereditary diseases, cancer,
tumours, deafness, X-chromosome-binding mental retardation and
lissencephalous disease. hBex is also useful for screening mimics,
agonists, or inhibitors, and in peptide fingerprinting identification.
hBex coding sequence can be used as primers or probes, or in producing
gene chips or microarrays.

XX SQ Sequence 792 BP; 214 A; 172 C; 219 G; 187 T; 0 other;

alignment_scores:
Quality: 184.00 Length: 131
Ratio: 2.453 Gaps: 5
Percent Similarity: 57.252 Percent Identity: 36.641

alignment_block:

US-09-327-750D-12 x AAH45143

Align seg 1/1 to: AAH45143 from: 1 to: 792

1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17

|||||.....
212 ATGGAATGCCAACCAAGAAATGAAGAAAGAGCAAGTTCCTAATAA 261

17 nGlyGluGlu...AspArgProValGlyGlyGlyGlyHisGlnProA 33

|||||.....
262 AGGGAGCCCTTGCCCTCCCTTTGGATGCTGTGAATACTGTGCTGA 311

33 laGlyAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisHisArg 49

|||||

312 GAGGAAAT.....CGTAGG 325

THIS PAGE BLANK (USPTO)

OM of: US-09-327-750D-12 to: GenEmbl:* out_format : pfs
Date: Mar 11, 2002 3:33 PM
About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2.1/uspto.spool/US09327750/runat_11032002.101154.20319/app_query.fasta_1.1472
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -MODALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -DOCS=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09327750@cgn2.1.8673
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-12
Query length: 124
Database: GenEmbl:
Database sequences: 1472140
Database length: -34134837
Search time (sec): 4557.230000

score_list:

Sequence	Strd	Orig	zScore	Escore	Len	Documentation
gb_ro:AF187066	+	688.00	1064.22	5.3e-51	700	AF187066 Mus musculus p75NTR-as
gb_ro:AF097440	+	688.00	1062.65	6.5e-51	854	AF097440 Mus musculus brain exp
gb_ro:AF187065	+	643.00	997.06	2.9e-47	519	AF187065 Rattus norvegicus p75NTR-a
gb_pr:BC003190	+	576.50	890.99	2.4e-41	793	BC003190 Homo sapiens, p75NTR-a
gb_pr:HUMOGC	+	576.50	890.07	2.7e-41	891	M38188 Human unknown protein fr
gb_pr:AF187064	+	576.50	890.07	2.7e-41	891	AF187064 Homo sapiens p75NTR-as
gb_sts:G72708	-	411.00	640.31	2.2e-27	421	G72708 MARC 4953-4954-991939031
gb_pr:HSV315F8	+	282.50	404.88	2.8e-14	45678	Z70719 Human DNA sequence frd
gb_pr:AK000959	+	284.50	405.55	2.6e-14	1229	AK000959 Homo sapiens cDNA FLJ
gb_pat:AX100231	+	284.50	404.73	2.9e-14	1364	AX100231 Sequence 32 from Pat
gb_pr:HS635G19	+	264.50	373.75	1.5e-12	69648	AL035494 Human DNA sequence f
gb_pr:HSY14B7	+	240.50	333.97	2.5e-10	98274	Z99755 Human DNA sequence frd
gb_sts:G35294	-	235.50	368.21	3.1e-12	477	G35294 human SFS SHGC-37409, se
gb_pr:AF183416	+	191.00	297.12	2.8e-08	642	AF183416 Homo sapiens ovarian g
gb_pr:AF220189	+	191.00	295.12	3.7e-08	828	AF220189 Homo sapiens uncharact
gb_pr:HS198P4	+	191.00	265.47	1.6e-06	35714	AL008708 Human DNA sequence f
gb_pr:AF237783	+	184.00	284.66	1.4e-07	791	AF237783 Homo sapiens brain exp
gb_ro:AF097439	+	172.00	266.19	1.5e-06	785	AF097439 Mus musculus brain exp
gb_pr:AF251053	+	171.50	265.84	1.6e-06	744	AF251053 Homo sapiens X-linked
gb_pat:AX078272	+	171.50	264.35	1.9e-06	898	AX078272 Sequence 76 from Pat
gb_pr:HSV870H8	+	171.50	236.38	6.9e-05	3321	Z70233 Human DNA sequence frd
gb_pr:AL133348	+	171.50	234.34	8.9e-05	40584	AL133348 Homo sapiens DNA sequence
gb_pr:HSU08B1	+	171.50	234.25	9.0e-05	41029	AL022169 Homo sapiens chromos
gb_ro:AF097438	+	156.00	240.98	3.8e-05	835	AF097438 Mus musculus brain exp
gb_ro:AF051347	+	156.00	240.95	3.8e-05	838	AF051347 Mus musculus REX-3 mRN
gb_ro:AF097437	+	156.00	233.11	0.0001	2269	AF097437 Mus musculus Bex1 pr
gb_pr:HSU10564	+	132.50	204.36	0.0042	43952	Z92846 Human DNA sequence frd
gb_htg:AL357046	+	138.00	172.06	0.2626	154577	AL357046 Homo sapiens chrom
gb_htg:AC009282	+	138.00	172.04	0.2633	154599	AC009282 Homo sapiens clone
gb_pr:AL353305	+	138.00	170.63	0.3153	185257	AL353305 Human DNA sequence
gb_sts:G24641	-	122.50	193.21	0.0174	504	G24641 human SFS WI-11354, sequ
gb_vi:AF181716	-	112.00	166.40	0.5425	1933	AF181716 Cynomolgus Epstein-Ba
gb_pr:HS112K5	+	106.50	126.08	95.54	109891	Z85987 Human DNA sequence fr
gb_htg:AL591766	+	106.00	121.35	175.26	181658	AL591766 Homo sapiens chrom
gb_htg:PFMAL4P3	+	104.00	113.78	462.96	321003	AL035476 Plasmodium falcipar
gb_htg:AC091781	+	103.00	117.62	282.95	162045	AC091781 Trypanosoma brucei
gb_in:LCU75355	+	101.00	145.11	8.32	3336	U75355 Lucilia cuprina ecyste
gb_vi:CH045963	-	100.50	148.62	5.30	1936	U45963 Cerothricine herpesvi
gb_htg:AC079989	+	100.50	112.68	532.71	185650	AC079989 Rattus norvegicus d
gb_htg:AC079378	+	100.50	112.58	539.96	188150	AC079378 Rattus norvegicus d

gb_ba:AE004899 - 100.00 134.31 33.27 10807 ! AE004899 Pseudomonas aerug
gb_pl:AB016876 + 100.00 123.16 138.96 4499 ! AB016876 Arabidopsis thali
gb_pl:NCB1D1 + 99.00 117.39 291.18 76072 ! AL355927 Neurospora crassa
gb_htg:AC011771 + 99.00 111.60 612.42 158807 ! AC011771 Homo sapiens chr
gb_pl:SCFPSIG + 97.00 140.33 15.37 2794 ! X54157 S.cerevisiae FPS1 (

seq_name: gb_ro:AF187066

seq_documentation_block:

LOCUS AF187066 700 bp mRNA ROD
DEFINITION Mus musculus p75NTR-associated cell death executor (Nade) mRNA,
complete cds.

ACCESSION AF187066

VERSION AF187066.1 GI:8452897

KEYWORDS

SOURCE house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 700)

REFERENCE

AUTHORS

Mukai,J., Hachiya,T., Shoji-Hoshino,S., Kimura,M.T., Nadano,D.,

Suvanto,P., Hanaoka,T., Li,Y., Irie,S., Greene,L.A. and Sato,T.A.

NADE, a p75NTR-associated cell death executor, is involved in

signal transduction mediated by the common neurotrophin receptor

p75NTR

J. Biol. Chem. 275 (23), 17566-17570 (2000)

MEDLINE 20298829

REFERENCE 2 (bases 1 to 700)

AUTHORS

Hanaoka,T., Li,Y., Irie,S. and Sato,T.

Direct Submission

Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia

University, 630 West 168th St., P&S 11-451, New York, NY 10032, USA

TITLE

JOURNAL

FEATURES

Source

1. .700

/organism="Mus musculus"

/strain="BALB/c"

/db_xref="taxon:10090"

1. .700

/gene="Nade"

177. .551

/gene="Nade"

/note="Involved in the common neurotrophin receptor

p75NTR-mediated signal transduction; NADE"

/codon_start=1

/product="p75NTR-associated cell death executor"

/protein_id="AAF75131.1"

/db_xref="GI:8452898"

/translation="NANYHQNEEMEQPLOGQEDRPVGGEGHQPAAANNNNNNHNNH

HNHRRGQARRLAPNFRWAIPIRNMNDGLGGPGDMEFMEMREIRKRLRELQRLNC

LRILMGLSNHHDHDEFLMP"

BASE COUNT 178 a 187 c 203 g 132 t

ORIGIN

alignment_scores:

Quality: 688.00 Length: 124

Ratio: 5.548 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 98.387

alignment_block:

US-09-327-750D-12 x AF187066

Align seg 1/1 to: AF187066 from: 1 to: 700

1 MetalaAsnValHisGlnGluAsnGluGluMetGlnProLeuGlnAs 17

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

177 ATGGCCAAATGTCACCAAGGAAAGAGATGGAGCAGCCCTGCAGAA 226

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

17 nGlyGluGluAspArgProValGlyGlyGlyGlyGlyHisGlnProAlaG 34

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

227 TGGACAGGAAGACCGCCCTGTGGGAGGAGGTGAGGCCAGCCTGCTG 276

34 lYAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisAsnHisArgArg 50
:|||||
277 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 326
51 GlyGlnAlaArgArgLeuAlaProAsnProAsnPheArgTrpAlaIleProAsnAr 67
|||||
327 GCCCAGGCTCGCCGACTTGCCTTAACCTTCCGATGGCCATTCCTCCACAG 376
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
|||||
377 GCAGATGAATGACGGGTGGGTGGAGATGGAGATGATGGAATGTTC 426
84 etGluGluMetArgGluLeuArgArgLysLeuArgGluLeuGlnLeuArg 100
|||||
427 TGGAGGAGATGAGAGATCCGAGAAAGCTTAGGGAGCTACACCTGAGA 476
101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisAspHisH 117
|||||
477 AATTGTCTACGCATCCTTATGGGGAGCTGTCTAACCCACACGATCACCA 526
117 sAspGluPheCysLeuMetPro 124
:|||||
527 TGATGAATTCGCTTATGCCT 548

seq_name: gb_ro:AF097440

seq_documentation_block: 854 bp mRNA ROD 13-APR-1999
LOCUS AF097440 Mus musculus brain expressed X-linked protein 3 (Bex3) mRNA,
DEFINITION complete cds.
ACCESSION AF097440
VERSION AF097440.1 GI:4580593
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 854)
AUTHORS Brown,A.L. and Kay,G.F.
TITLE Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
JOURNAL Hum. Mol. Genet. 8 (4), 611-619 (1999)
MEDLINE 99172070
REMARK Erratum:[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
REFERENCE 2 (bases 1 to 854)
AUTHORS Brown,A.L. and Kay,G.F.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
FEATURES
Location/Qualifiers
1..854
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/db_xref="dbEST:AA272375"
/chromosome="X"
/map="near Plp"
/cissue_type="pooled organs"
1..854
/gene="Bex3"
172..546
/gene="Bex3"
/codon_start=1
/product="brain expressed X-linked protein 3"
/protein_id="AAD24431.1"
/db_xref="GI:4580594"
/translation="MANVHQNEEMEQPLQGOEDRPVGGEGHQPAAANNHNNHNNH
HNHRRGOARLAPNFRWAIPIRQNDGLGGDDMEMEMREIRKRLRELQRLNC
LRILGELSNHHHDHDEFLMP"

BASE COUNT 237 a 212 c 228 g 177 t
ORIGIN

alignment_scores: Quality: 688.00 Length: 124
Ratio: 5.548 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.387
alignment_block:
US-09-327-750D-12 x AF097440 ..
Align seg 1/1 to: AF097440 from: 1 to: 854
1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
|||||
172 ATGCCCAATGTCCACAGAAAGAAAGAGATGGAGCGCCCTGCAGAA 221
17 nGlyGluGluAspArgProValGlyGlyGlyGlyGlyHisGlnProAlaG 34
|||||
222 TGGACAGAAAGACCGCCCTGTGGGAGGAGGTGAGGGCCACGAGCTGCTG 271
34 lYAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisAsnHisArgArg 50
:|||||
272 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 321
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 67
|||||
322 GCCCAGGCTCGCCGACTTGCCTTAACCTTCCGATGGCCATTCCTCCACAG 371
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
|||||
372 GCAGATGAATGACGGGTGGGTGGAGATGGAGATGATGGAATGTTC 421
84 etGluGluMetArgGluLeuArgArgLysLeuArgGluLeuGlnLeuArg 100
|||||
422 TGGAGGAGATGAGAGATCCGAGAAAGCTTAGGGAGCTACAGCTGAGA 471
101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisAspHisH 117
|||||
472 AATTGTCTACGCATCCTTATGGGGAGCTGTCTAACCCACACGATCACCA 521
117 sAspGluPheCysLeuMetPro 124
:|||||
522 TGATGAATTCGCTTATGCCT 543

seq_name: gb_ro:AF187065

seq_documentation_block: 519 bp mRNA ROD 11-JUN-2000
LOCUS AF187065 Rattus norvegicus p75NTR-associated cell death executor (Nade)
DEFINITION mRNA, complete cds.
ACCESSION AF187065
VERSION AF187065.1 GI:8452895
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 519)
AUTHORS Mukai,J., Hachiya,T., Shoji-Hoshino,S., Kimura,M.T., Nadano,D.,
Suvanto,P., Hanaoka,T., Li,Y., Irie,S., Greene,L.A. and Sato,T.A.
TITLE Nade, a p75NTR-associated cell death executor, is involved in
signal transduction mediated by the common neurotrophin receptor
p75NTR
JOURNAL J. Biol. Chem. 275 (23), 17566-17570 (2000)
MEDLINE 20298829
REFERENCE 2 (bases 1 to 519)
AUTHORS Mukai,J., Hachiya,T., Hoshino,S., Kimura,M., Nadano,D.,
Hanaoka,T., Li,Y., Irie,S. and Sato,T.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia
University, 630 West 168th St., P&S 11-451, New York, NY 10032, USA
FEATURES
Location/Qualifiers
1..519
source


```
34  lyAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArgArg 50
|||||
296  GAAAT.....CGACGG 306

51  GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 67
|||||
307  GGACAGGCTCGCCGACTTGGCCCTTAATTTTCGATGGGCCATACCAANTAG 356

67  gGlnMetAsnAspGlyLeuGlyArgGlyAspGlyAspMetGluMetPheM 84
|||||
357  GCAGATCAATGATGGGTGGGTGGAGATGGAGATGATGATGAAATATTC 405

84  etGluGluMetArgGluIleArgArgLysLeuArgGluLeuGlnLeuArg 100
|||||
407  TGGAGGAGATGAGAGAAATCAGAAGAACTTGGGAGCTGCAGTTGAGG 456

101  AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisH 117
|||||
457  AATTGTCTGCGTATCCTTATGGGGAGCTCTCTAATCACCATGACCATCA 506

117  sAspGluPheCysLeuMetPro 124
|||||
507  TGATGAATTTTGCCTTATGCCT 528

seq_name: gb_pr:HUMOGC

seq_documentation_block:
LOCUS HUMOGC 891 bp mRNA 07-MAR-1995
DEFINITION Human unknown protein from clone pHR74 mRNA, complete cds.
ACCESSION M38188 X56942
VERSION M38188.1 GI:189378
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Rapp,G., Freudenstein,J., Klaudiny,J., Mucha,J., Wempe,F.,
Zimmer,M. and Scheit,K.H.
TITLE Characterization of three abundant mRNAs from human ovarian
granulosa cells
JOURNAL DNA Cell Biol. 9 (7), 479-485 (1990)
MEDLINE 91025550
COMMENT [Draft entry and computer-readable sequence for [DNA 9, 479-485
(1990)]] kindly submitted
by K.H.Scheit, 27-AUG-1990.
FEATURES
source
Location/Qualifiers
1..891
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="pHR74"
/cell_type="granulosa"
/tissue_type="ovary"
<1..891
/note="protein of unknown function"
312..547
/note="protein of unknown function"
/codon_start=1
/protein_id="AAA63232.1"
/db_xref="GI:189379"
/translation="MANIHOENEEMOPMONGEEDRPLGGEGHQPAGNRREGQARRLA
PNFRWAINRINDMGDDGDDMEIFMEEMREIRKRLRELQRLNCLRLMGELSNHHD
HHDEFCLMP"
361..534
/note="protein of unknown function"
/codon_start=1
/protein_id="AAA63233.1"
/db_xref="GI:189380"
/translation="MERKALTALVEEVKATSLQEIQDRIADLPILFDGPPYPIGRSMGW
VEMEMIKYSWRR"
BASE COUNT 251 a 182 c 224 g 234 t
ORIGIN
```

```
alignment_scores:
Quality: 576.50 Length: 124
Ratio: 5.194 Gaps: 1
Percent Similarity: 89.516 Percent Identity: 84.677

alignment_block:
US-09-327-750D-12 x HUMOGC ..
Align seg 1/1 to: HUMOGC from: 1 to: 891

1 MetAlaAsnValHisGlnGluAsnGluGluMetGluGlnProLeuGlnAs 17
|||||
312 ATGGCAAAATATTACCAGGAAACGAGAGATGGAGCGCTATGCAGAA 361

17 nGlyGluGluAspArgProValGlyGlyGlyGlyGlyHisGlnProAlaG 34
|||||
362 TGCAGAGGAGACCGCCCTTTGGGAGGAGGTGAAGGCCACCGCTCCAG 411

34 lyAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisArgArg 50
|||||
412 GAAAT.....CGACGG 422

51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 67
|||||
423 GGACAGGCTCGCCGACTTGGCCCTTAATTTTCGATGGGCCATACCAANTAG 472

67 gGlnMetAsnAspGlyLeuGlyArgGlyAspGlyAspMetGluMetPheM 84
|||||
473 GCAGATCAATGATGGGTGGGTGGAGATGGAGATGATGGAATATTC 522

84 etGluGluMetArgGluIleArgArgLysLeuArgGluLeuGlnLeuArg 100
|||||
523 TGGAGGAGATGAGAGAAATCAGAAGAACTTGGGAGCTGCAGTTGAGG 572

101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisH 117
|||||
573 AATTGTCTGCGTATCCTTATGGGGAGCTCTCTAATCACCATGACCATCA 622

117 sAspGluPheCysLeuMetPro 124
|||||
623 TGATGAATTTTGCCTTATGCCT 644

seq_name: gb_pr:AF187064

seq_documentation_block:
LOCUS AF187064 891 bp mRNA PRI 11-JUN-2000
DEFINITION Homo sapiens p75NTR-associated cell death executor (NADE) mRNA,
complete cds.
ACCESSION AF187064
VERSION AF187064.1 GI:8452893
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Mukai,J., Hachiya,T., Hoshino,S., Kimura,M., Nadano,D.,
Suvanto,P., Hanaoka,T., Li,Y., Irie,S., Greene,L.A. and Sato,T.A.
TITLE NADE, a p75NTR-associated cell death executor, is involved in
signal transduction mediated by the common neurotrophin receptor
p75NTR
JOURNAL J. Biol. Chem. 275 (23), 17566-17570 (2000)
MEDLINE 20298829
REFERENCE
AUTHORS Mukai,J., Hachiya,T., Hoshino,S., Kimura,M., Nadano,D.,
Suvanto,P.,
Hanaoka,T., Li,Y., Irie,S. and Sato,T.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia
University, 630 West 168th St., P&S 11-451, New York, NY 10032, USA
FEATURES
Location/Qualifiers
1..891
source
```



```
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. .891
/gene="NADE"
312. .647
/gene="NADE"
/feature="NADE"
/feature="Involved in the common neurotrophin receptor
p75NTR-mediated signal transduction"
/codon_start=1
/product="p75NTR-associated cell death executor"
/protein_id="AAF75129.1"
/db_xref="GI:8452894"
/translation="WANIHQENEEMEQPNQGEEDRPLGGEGHQPAGNRRQARRLA
PNFRWAIPLPNDMGDDMEIFMEEMREIRKLRELQLRNLRLMGLSNHHD
HIDEFLAP"
BASE COUNT 251 a 182 c 224 g 234 t
ORIGIN
1 MetAlaSnValHisGlnGlnAsnGluMetGluGlnProLeuGlnAs 17
312 ATGGCAATATATTCACAGGAAACAGAGAGATGGAGCGCTATGCAGAA 361
17 nGlyGluGluAspArgProValGlyGlyGlyGlyGlyHisGlnProAlaG 34
362 TGGAGAGAGAGAGCGCCCTTGGGAGGAGGTGAAGGCCACCGCTCGAG 411
34 lyAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnArg 50
412 GAAAT.....CGACGG 422
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaLeuProAsnAr 67
423 GGACAGGCTCGCCGACTTGGCCCTTAATTTTCGATGGGCCATACCAATAG 472
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
473 GCATCAATGATGGGATGGGTGGAGATGGAGATGATGGAATATTC 522
84 etGluGluMetArgGluLeuArgArgLysLeuArgGluLeuGlnLeuArg 100
523 TGGAGGAGATGAGAGAAATCAGAGAGAACTTAGGGAGCTGCAGTTCAGG 572
101 AsnCysLeuArgIleLeuMetGlyLeuSerAsnHisAspHisH 117
573 AATTTCTCGTATCTCTATGGGGAGCTCTCTAATCACCATTGACCATCA 622
117 sAspGluPheCysLeuMetPro 124
623 TGATGAATTTGCCTTATGCCT 644
seq_name: gb_sts:G72708
seq documentation block: 421 bp DNA STS 08-AUG-2001
LOCUS G72708
DEFINITION MARC 4953-4954:991939031:1 SCF - porcine spleen Sus scrofa STS
genomic, sequence tagged site.
ACCESSION G72708
VERSION G72708.1 GI:15146738
KEYWORDS STS.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
REFERENCE
AUTHORS Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 421)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L. and
Keeler,J.W.
TITLE Single nucleotide polymorphism (SNP) discovery in expressed porcine
genes
JOURNAL Unpublished (2001)
COMMENT
CONTACT: Freking BA
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4278
Fax: 402 762 4173
Email: freking@email.marc.usda.gov
Primer A: GCAATGGTGCAACTCTACT
Primer B: CCAGCAGCAATAGACG
STS size: 500
PCR profile:
Hotstart: 95 degrees for 15 minutes
Denature: 95 degrees for 30 seconds
Anneal: 56 degrees
Extension: 68 degrees for 2 minutes
Cycles: 32 to 45
Protocol:
Template: 50-200 ng genomic DNA
Primer: each 20 pmoles
dNTPs: each 88 uM
Taq Polymerase: 0.25 units (Qiagen HotStar)
Buffer: Commercially supplied Qiagen HotStar buffer
The STS is derived from PCR amplicons generated from genomic DNA,
sequenced from each end using the amplification primers. The
sequence does not necessarily represent the entire amplicon.
Sequence derived from PolyPhred was trimmed from each end of each
unique contig until five consecutive bases exceeded a quality score
threshold of 20, and the next 10 bases averaged a quality score of
20 or greater. Amplicon size was estimated by agarose gel
electrophoresis.
FEATURES
Location/Qualifiers
1. .421
/organism="Sus scrofa"
/strain="white composite, duroc, meishan, minzhu,
fengjing, crossbreds"
/db_xref="taxon:9823"
/sex="male and female"
/clone_lib="SCF - porcine spleen"
/dev_stage="adult"
/note="Organ: spleen"
<1. >421
BASE COUNT 130 a 101 c 86 g 102 t 2 others
ORIGIN
alignment_scores:
Quality: 411.00 Length: 77
Ratio: 5.338 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 96.104
alignment_block:
US-09-327-750D-12 x G72708/rev ..
Align seg 1/1 to reverse of: G72708 from: 1 to: 421
48 HisArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIle 64
:::|||||
421 AATAGACGGGACAAAGCTGCGGACTTGGCCCTAATTTCCGATGGCCAT 372
64 eProAsnArgGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetG 81
|||||
371 ACCCAATAGCAGATCAATGATGGATGGGTGGAGATGGAGATGATATGG 322
```

81 luMetPheMetGluGluMetArgGluIleArgArgLysLeuArgGluLeu 97
 |||||
 321 AAATGTTATGAGAGATGAGAAATCAGAGAAACTTAGGAGCTG 272
 |||||
 98 GlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisH 114
 |||||
 271 CAGTTGAGGAATGCTGCTATCCTTATGGGGAGCTCTCTAATCACCA 222
 |||||

114 sasphishisAspGluPheCysLeuMetPro 124
 |||||
 221 TGACCATCATGATGAATTTGCCTATGCCT 191
 |||||

seq_name: gb_pr:HSV351F8

seq_documentation_block:

LOCUS HSV351F8 45678 bp DNA PRI 23-NOV-1999
 DEFINITION Human DNA sequence from cosmid V351F8, between markers DXS366 and
 DXS87 on chromosome X contains ESTs.

ACCESSION 270719

VERSION 270719.1 GI:1261915

KEYWORDS X.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 45678)
 AUTHORS Whiteley,M.
 TITLE Direct Submission

JOURNAL Submitted (09-APR-1995) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1RQ, UK. E-mail enquiries: humquerry@sanger.ac.uk
 COMMENT IMPORTANT: This sequence is the entire insert of clone V351F8. The
 true left end of clone V351F8 is at 1 in this sequence. The true
 right end of clone V351F8 is at 45678.
 V351F8 is from the human chromosome X-specific cosmid library.

FEATURES

source

1..45678

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="X"

/map="X"

/clone="GHC-351F8"

/clone_lib="SCCV"

1..70

repeat_region

/partial

/note="Alu repeat: matches 80..1 of consensus"

642..930

repeat_region

/note="Alu repeat: matches 1..308 of consensus"

1406..5179

repeat_region

/note="L1 element fragment"

5243..5276

repeat_region

/note="17 copies of 2 mer 82 % conserved"

5279..5567

repeat_region

/partial

/note="Alu repeat: matches 304..1 of consensus"

9975..10369

repeat_region

/note="MSTA element fragment"

10399..12017

repeat_region

/note="MSTAR element fragment"

11318..11978

repeat_region

/note="THE1BR element fragment"

11715..11846

repeat_region

/note="MLT1R element fragment"

12074..12170

repeat_region

/note="MSTA element fragment"

12227..12486

repeat_region

/note="MSTA element fragment"

12437..12502

repeat_region

/note="MSTC element fragment"

12756..13059

repeat_region

/partial

/note="Alu repeat: matches 308..1 of consensus"

14478..14650

repeat_region

/note="THE1B element fragment"

repeat_region 14491..14536
 /note="MSTA element fragment"
 repeat_region 14715..14831
 /note="THE1B element fragment"
 repeat_region 15099..15533
 /note="L1 element fragment"
 repeat_region 15534..15626
 /note="31 copies of 3 mer 85 % conserved"
 repeat_region 15566..15619
 /note="3 copies of 18 mer 98 % conserved"
 repeat_region 16352..16436
 /note="L1 element fragment"
 repeat_region 16728..16763
 /note="9 copies of 4 mer 94 % conserved"
 repeat_region 16764..17034
 /partial
 repeat_region 19420..19449
 /note="Alu repeat: matches 308..1 of consensus"
 repeat_region 19420..19449
 /note="15 copies of 2 mer 87 % conserved"
 misc_feature complement(19537..20013)
 /note="match: 3' EST N51315 clone 283089"
 repeat_region 20071..20114
 /note="22 copies of 2 mer 98 % conserved"
 repeat_region 20073..20112
 /note="10 copies of 4 mer 100 % conserved"
 repeat_region 20073..20126
 /note="3 copies of 18 mer 87 % conserved"
 repeat_region 23088..23300
 /note="L1 element fragment"
 repeat_region 23311..23490
 /note="MLT1A element fragment"
 repeat_region 23318..23489
 /note="MLT1B element fragment"
 repeat_region 23362..23491
 /note="MLT1C element fragment"
 repeat_region 23507..24407
 /note="L1 element fragment"
 repeat_region 24376..24462
 /note="MSTC element fragment"
 repeat_region 24379..24491
 /note="MSTA element fragment"
 repeat_region 24688..24738
 /note="MSTA element fragment"
 repeat_region 24810..25016
 /note="L1 element fragment"
 repeat_region 25254..25421
 /note="L1 element fragment"
 repeat_region 25439..25696
 /note="MER25 element fragment"
 repeat_region 26284..26343
 /note="MLT1D element fragment"
 repeat_region 26428..26491
 /note="MLT1D element fragment"
 repeat_region 26465..26518
 /note="MLT1E element fragment"
 repeat_region 27446..27667
 /note="L1 element fragment"
 repeat_region 28954..29247
 /partial
 repeat_region 30160..30325
 /note="Alu repeat: matches 308..1 of consensus"
 repeat_region 30160..30325
 /note="2 copies of 83 mer 98 % conserved"
 misc_feature 32588..32875
 /note="match: 5' EST H68599 clone 239077"
 misc_feature complement(32825..33230)
 /note="match: 3' EST H68239 clone 289077"
 repeat_region 35551..35802
 /partial
 repeat_region 35804..35839
 /note="Alu repeat: matches 1..260 of consensus"
 repeat_region 35804..35839
 /note="18 copies of 2 mer 83 % conserved"
 repeat_region 38625..38705
 /note="MLT1R element fragment"


```

repeat_region 15994..16040
repeat_region 16116..16184
repeat_region 16897..17005
repeat_region 17618..17700
repeat_region 17748..17912
repeat_region 18657..18871
repeat_region 18927..19230
repeat_region 19492..19559
repeat_region 19560..19821
repeat_region 19822..19959
repeat_region 19987..20286
repeat_region 20316..20512
repeat_region 20513..20741
repeat_region 20769..21074
repeat_region 21085..21207
repeat_region 21212..21391
repeat_region 21392..21471
repeat_region 22146..22441
repeat_region 22526..22610
repeat_region 22940..23204
repeat_region 23205..23514
repeat_region 23919..24194
repeat_region 24496..24590
repeat_region 24871..24916
repeat_region 25378..25555
repeat_region 25673..25910
repeat_region 26722..26798
repeat_region 27040..27107
repeat_region 27108..27374
repeat_region 27375..27471
repeat_region 27651..27945
repeat_region 28235..28876
repeat_region 28975..29097
repeat_region 29173..29196
repeat_region 29832..30260
misc_feature 30392..30597
repeat_region 30845..31010

```

```

/note="MER91B repeat: matches 2..162 of consensus"
31056..31094
/note="MIR repeat: matches 218..256 of consensus"
31728..31791
/note="16 copies 4 mer t9tg 76 conserved"
31858..31952
/note="MIR repeat: matches 48..153 of consensus"
complement(33437..33589)
/note="match: STS: Em:H5731"
33543..33633
/note="MIR repeat: matches 84..165 of consensus"
33785..33966
/note="MER91A repeat: matches 1..185 of consensus"

```

```

alignment_scores:
  Quality: 240.50      Length: 67
  Ratio: 4.076        Gaps: 2
  Percent Similarity: 88.060  Percent Identity: 77.612

```

```
alignment_block:
```

```
US-09-327-750D-12 x HS714B7
```

```
Align seg 1/1 to: HS714B7 from: 1 to: 98274
```

```

56 LeuAlaProAsnPheArgTrpAlaIleProAsnArgGlnMetAsnAspG1 72
|||||
62887 TTGGCCCTTAATTTCAATGGCCATACCAATAGGAGTCAATGATGG 62936

72 yLeuGlyAspGlyAspAspMetGluMetPheMetGluGluMetArgG 89
|||||
62937 GATCAATAGAGTGGAGATGATGGAATGTTTCATGGAGGAGATGAG 62986

89 luileArgArg.LysLeuArgGluLeuGluLeuArgAsnCysLeuArg11 105
|||||
62987 GAATCGAGGAGAGAAT...AAGGAGCTACAATGAGGAATTTCTGTGTAT 63033

105 eLeuMetGlyGluLeuSerAsnHisAspHisAspGluPheCys 121
|||||
63034 CCTATGGGAAG.CTGTGTAATCCCATGACCATCATGATGAATTTGCC 63081

```

```
seq_name: gb_sts:G35294
```

```
seq_documentation_block:
LOCUS G35294 477 bp DNA STS 02-OCT-1997
DEFINITION human STS SHGC-37409, sequence tagged site.
```

```

ACCESSION G35294
VERSION G35294.1 GI:2459462
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 477)
AUTHORS Myers, R.M.
JOURNAL Unpublished (1997)
COMMENT

```

```

Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu

```

```

Primer A: AACATCTTTCATGAAAGTTGATG
Primer B: CTTTGTGGCATCTTCTGCAA
STS size: 106
PCR Profile:

```

```

Initial incubation: 95 degrees C for 10 minutes
Denaturation: 94 degrees C for 15 seconds
Annealing: 60 degrees C for 30 seconds
Polymerization: 72 degrees C for 23 seconds

```


94 LeuArgGluLeuGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLe 110
||||| ||||| :::: ||||| :::
||||| ||||| :::: ||||| :::

THIS PAGE BLANK (USPTO)

OM of: US-09-327-750D-12 to: EST:* out_format : pfs

Date: Mar 11, 2002 2:16 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=Cgnt2.1/USPTO.spool/US09327750/runat_11032002_101153_20308/app_query.fasta_1.1472
-DB=EST_QPWT-fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000
-DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTPWT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=2000000000 -USER=US09327750_ECGN1_1.5654
-NCPU=6 -ICPU=3 -IONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-12

Query length: 124

Database: EST:

Database sequences: 11351937

Database length: 1077921985

Search time (sec): 4085.940000

score_list:

Sequence	Strd	Orig	zScore	Escore	Len	Documentation
gb_est2:W46041	+	688.00	1165.70	9.1e-56	492	!
gb_est1:AI006575	+	688.00	1164.93	1.0e-55	538	!
gb_est1:BE334866	+	688.00	1164.87	1.0e-55	542	!
gb_est1:AA215070	+	688.00	1164.63	1.0e-55	557	!
gb_est1:BE334877	+	688.00	1164.21	1.1e-55	595	!
gb_est2:W64711	+	688.00	1164.13	1.1e-55	590	!
gb_est2:RG088461	+	688.00	1164.00	1.1e-55	599	!
gb_est1:AA272375	+	688.00	1163.82	1.2e-55	612	!
gb_est2:W81757	+	688.00	1163.76	1.2e-55	616	!
gb_est1:W476468	+	688.00	1163.60	1.2e-55	628	!
gb_hic:AK004531	+	688.00	1161.41	1.6e-55	809	!
gb_est1:AA253897	+	687.00	1164.44	1.1e-55	468	!
gb_hic:AK010500	+	683.00	1152.56	4.9e-55	845	!
gb_est1:AI152323	+	681.00	1155.00	3.6e-55	430	!
gb_est1:AW908751	+	679.00	1150.26	6.6e-55	503	!
gb_est2:BF681983	+	675.00	1149.08	7.7e-55	577	!
gb_est1:AA268306	+	675.00	1142.30	1.8e-54	577	!
gb_est1:AI118980	+	673.00	1141.38	2.1e-54	642	!
gb_est2:BF012441	+	673.00	1141.45	2.0e-54	430	!
gb_est1:AF636864	+	673.00	1139.99	2.5e-54	509	!
gb_est2:BF608455	+	673.00	1136.23	4.0e-54	787	!
gb_est1:AA475682	+	667.00	1131.70	7.2e-54	410	!
gb_est1:AA044548	+	661.00	1120.01	3.2e-53	489	!
gb_est1:AA466788	+	661.00	1119.15	3.6e-53	540	!
gb_est2:BI407164	+	661.00	1115.58	5.7e-53	817	!
gb_est2:W98871	+	659.00	1115.83	5.5e-53	536	!
gb_est1:AA667520	+	658.00	1115.09	6.0e-53	480	!
gb_hic:AK003294	+	657.00	1108.71	1.4e-52	826	!
gb_est1:AA004191	+	654.00	1107.26	1.6e-52	542	!
gb_est1:AA050176	+	652.00	1103.86	2.5e-52	543	!
gb_est1:AA606578	+	650.00	1100.63	3.8e-52	533	!
gb_est1:AW476461	+	645.00	1091.93	1.2e-51	547	!
gb_est1:AA049613	+	642.00	1087.52	2.1e-51	506	!
gb_est1:BE198421	+	640.00	1084.98	2.9e-51	459	!
gb_est1:W57617	+	640.00	1082.62	3.9e-51	603	!
gb_est2:BI082462	+	639.00	1077.93	7.1e-51	853	!
gb_est1:AA214902	+	637.00	1078.60	6.5e-51	533	!
gb_est1:AA259786	+	634.00	1074.32	1.1e-50	486	!
gb_est1:AA110311	+	634.00	1074.07	1.2e-50	500	!
gb_est1:AA068855	+	631.00	1068.62	2.3e-50	522	!

gb_est2:BI281721 - 628.00 1061.48 5.8e-50 662 ! BI281721 UI-R-CT0s-cav-f-10-
gb_est2:BF168928 + 621.50 1048.73 3.0e-49 810 ! BF168928 60177532F1 NCI_CGA
gb_est1:AA214909 + 618.00 1046.43 4.0e-49 532 ! AA214909 mu76f05.rl Stratagene
gb_est1:AI183075 + 615.00 1042.69 6.5e-49 455 ! AI183075 ub93e07.rl Soares_m
gb_est1:AI157489 + 611.00 1037.80 1.2e-48 366 ! AI157489 ue56e05.rl Soares_m

seq_name: gb_est2:W46041

seq_documentation_block:

LOCUS W46041 492 bp mRNA EST 23-MAY-1996

DEFINITION mc82h02.rl Soares mouse embryo NBME13.5 14.5 Mus musculus CDNA

clone IMAGE:355059 5' similar to gb:M38188 OVARIAN GRANULOSA CELL

13.0 KD PROTEIN HGR74 (HUMAN); mRNA sequence.

ACCESSION W46041

VERSION W46041.1 GI:1330757

KEYWORDS EST

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 492)

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

Waterston,R.

TITLE The WashU-HMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LML; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:226859

Seq primer: ETPrimer

High quality sequence stop: 489.

Location/Qualifiers

1..492

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:355059"

/clone_lib="Soares mouse embryo NBME13.5 14.5"

/sex="unknown"

/tissue_type="embryo"

/dev_stage="13.5-14.5dpc total fetus"

/lab_host="PH108"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site1: Not I; Site2: Eco RI; 1st strand CDNA

was primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAATGGCGCGGAAATTTTTTTTTTTTTTTTTTTT

T 3'], on equal amounts of mRNA from 2 13.5dpc and 2

14.5dpc embryos [total RNA provided by Minoru Ko, Wayne

State Univ., from 2]; double-stranded cDNA was ligated to

Eco RI adaptors (Pharmacia), digested with Not I and

cloned into the Not I and Eco RI sites of the modified

pT73 vector. Library went through one round of

normalization, and was constructed by Bento Soares and

M.Fatima Bonaldo."

BASE COUNT 133 a 133 c 125 g 101 t

ORIGIN

alignment_scores:

Quality: 688.00 Length: 124

Ratio: 5.548 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 98.387

alignment_block:

US-09-327-750D-12 x W46041

Align seg 1/1 to: W46041 from: 1 to: 492

```

1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
13 ATGGCCAAATGTCACCAAGAAACGAGAGATGGAGCAGCCCTGCAGAA 62
17 nGlyGluGluAspArgProValGlyGlyGlyGlyGlyHisGlnProAlaG 34
63 TGGACAGGAGACCGCCCTGTGGAGAGGTGAGGGCCACCAAGCCTGCTG 112
34 lyAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHis 50
113 CAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 162
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsn 67
163 GGCACAGGCTCGCGGACTTGCCTTAACCTCCGATGGGCCATTCACCA 212
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
213 GCAGATGATGACGGGTGGGTGGAGATGGAGATGATGGAAATGTTCA 262
84 etGluGluMetArgGluIleArgArgLysLeuArgGluGluLeuArg 100
263 TGGAGGAGATGAGAGATCGGAGAAAGCTTAGGAGCTACAGCTGAGA 312
101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisAsnHisAs 117
313 AATTGCTACGCATCTTTATGGGGAGCTGTCTAACCAACCACCATCAC 362
117 sAspGluPheCysLeuMetPro 124
363 TGATGAATTCGCTTATGCCT 384

```

seq_name: gb_estl:A1006575

```

seq_documentation_block: 538 bp mRNA EST 12-JUN-1998
LOCUS A1006575
DEFINITION uel506.y1 Sugano mouse embryo mewa Mus musculus cDNA clone
IMAGE:1480450 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION A1006575
VERSION A1006575
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 538)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilsson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

```

```

This clone is available royalty-free through LILN; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:928806
Seq primer: primer name ambiguous
High quality sequence stop: 433.
Location/Qualifiers
1..538
/organism="Mus musculus"

```

FEATURES

```

seq_name: gb_estl:BE334866
seq_documentation_block:
LOCUS BE334866
DEFINITION us90a11.y1 Soares mammary_gland_NWLMG Mus musculus cDNA clone
IMAGE:3325628 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION BE334866

```

```

/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1480450"
/clone_lib="Sugano mouse embryo mewa"
/dev_stage="embryo, 14 dpc"
/lab_host="DH10B"
/note="vector: pME18S-FL3; Site 1: DraIII (CACTGTGTG);
Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed
with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTTCTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[GTGTGGCCTACTGG], digested and cloned into distinct DraIII
sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site
CACCATGTG). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed by Dr. Sumio Sugano
(University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CTTCTGCTCTAAAGCTGCG and 3' end primer
CGACCTGCAGCTCGACCA."
BASE COUNT 154 a 140 c 162 g 82 t
ORIGIN

```

```

alignment_scores:
Quality: 688.00 Length: 124
Ratio: 5.548 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.387

```

alignment_block:

US-09-327-750D-12 x A1006575

Align seg 1/1 to: A1006575 from: 1 to: 538

```

1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
162 ATGGCCAAATGTCACCAAGAAACGAGAGATGGAGCAGCCCTGCAGAA 211
17 nGlyGluGluAspArgProValGlyGlyGlyGlyGlyHisGlnProAlaG 34
212 TGGACAGGAGACCGCCCTGTGGAGAGGTGAGGGCCACCAAGCCTGCTG 261
34 lyAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisAsnHis 50
262 CAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 311
51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsn 67
312 GGCACAGGCTCGCGGACTTGCCTTAACCTCCGATGGGCCATTCACCA 361
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
362 GCAGATCAATGACGGGTGGGTGGAGATGGAGATGATGGAAATGTTCA 411
84 etGluGluMetArgGluIleArgArgLysLeuArgGluGluLeuArg 100
412 TGGAGGAGATGAGAGATCGGAGAAAGCTTAGGAGCTACAGCTGAGA 461
101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisAsnHisAsn 117
462 AATTGCTACGCATCTTTATGGGGAGCTGTCTAACCAACCACCATCAC 511
117 sAspGluPheCysLeuMetPro 124
512 TGATGAATTCGCTTATGCCT 533

```

seq_name: gb_estl:BE334866

```

VERSION BE334866.1 GI:9208642
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota: Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1069792
Seq primer: -40RP from Gibco
High quality sequence stop: 465.
FEATURES
    source
        1..542
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /clone="IMAGE:3325628"
            /clone_lib="Soares_mammary_gland_NMLMG"
            /sex="female (lactating)"
            /tissue_type="mammary gland"
            /lab_host="DH10B"
            /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT7T3 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 156 a 141 c 164 g 81 t
ORIGIN

alignment_scores
    Quality: 688.00 Length: 124
    Ratio: 5.548 Gaps: 0
    Percent Similarity: 100.000 Percent Identity: 98.387

alignment_block
    US-09-327-750D-12 x BE334866 ..
    Align seg 1/1 to: BE334866 from: 1 to: 542

1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
166 ATGCCCAATGTCACACAGAAAGAGAGATGGAGCAGCCCTGCAGAA 215
17 nGlyGluAspArgProValGlyGlyGlyGlyGlyHisGlnProAlaG 34
216 TGGACAGAGAGACCGCCTCTGGGAGAGGTGAGGGCCACAGCCTGCTG 265
34 lYAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArg 50
266 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 315
51 GlyGlnAlaArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 67
316 GGCAGAGCTCCGACACTGCCCTTAACCTCCGATGGCCATCCCAACAG 365
67 gGlnMetAsnAspGlyLeuGlyGlyGlyGlyGlyGlyGlyGlyGly 84
366 GCAGATGAATGACGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGCA 415
84 etGluGluMetArgGluIleArgArgLysLeuArgGluLeuGlnLeuArg 100
416 TGGAGGAGATGAGAGATCCGGAGAAAGCTTAGGGAGCTACAGCTGAGA 465

101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisH 117
146 AATTGTCTACGCAATCTTATGGGGAGCTGTCTAACCCACACGATCACA 515

117 sAspGluPheCysLeuMetPro 124
516 TGATGAATTCCTGCCTTATGCCT 537

seq_name: gb_estl:AA215070

seq_documentation_block:
LOCUS AA215070 557 bp mRNA EST 03-FEB-1997
DEFINITION mu76f04.r1 Stratagene mouse embryonic carcinoma (#937317) Mus
musculus cDNA clone IMAGE:651487 5' similar to gb:M38188 OVARIAN
GRANULOSA CELL 13.0 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION AA215070
VERSION AA215070.1 GI:1814831
KEYWORDS EST:
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota: Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 557)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:397335
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 447.
FEATURES
    source
        1..557
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /clone="IMAGE:651487"
            /clone_lib="Stratagene mouse embryonic carcinoma (#937317)
"
            /tissue_type="carcinoma"
            /dev_stage="embryonic"
            /lab_host="SOLR (kanamycin resistant)"
            /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dt. p19 cell
line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT 159 a 145 c 165 g 88 t
ORIGIN

alignment_scores
    Quality: 688.00 Length: 124
    Ratio: 5.548 Gaps: 0
    Percent Similarity: 100.000 Percent Identity: 98.387

alignment_block
    US-09-327-750D-12 x AA215070 ..
    Align seg 1/1 to: AA215070 from: 1 to: 557

1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17

```

141 ATGGCAATGTCACCAAGAAACGAGAGATGGACCAACCCCTGCAGAA 190
 17 nGlyGluAspArgProValGlyGlyGluGlyHisGlnProAlaG 34
 191 TGCACAGGAAGACCGCCCTGTGGAGGAGGTGAGGGCCACCACTGCTG 240
 34 lYAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArg 50
 241 CAAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 290
 51 GlyGlnAlaArgLeuAlaProAsnPheArgTTPAlaIleProAsn 67
 291 GCCAGGCTCGCGGACTTCCCTTAACTTCCGATGGCCATTCCCAAC 340
 67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPhe 84
 341 GCAGATGAATGACGGTGTGGGTGGAGATGGAGATGATGGAATGTTCA 390
 84 etGluGluMetArgGluIleArgArgLysLeuArgGluLeuGlnLeu 100
 391 TGGAGGAGATGAGAGAGATCCGAGAAAGCTTAGGGAGCTACAGCTG 440
 101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisAsnHis 117
 441 AATTGCTACGATCTTATGGGGAGCTGTCTAACCAACCAAGATCAC 490
 117 sAspGluPheCysLeuMetPro 124
 491 TGATGAATTCGCTTATGCGCT 512

seq_name: gb_est1:BE334877

seq_documentation_block:
 LOCUS BE334877 585 bp mRNA EST 14-JUL-2000
 DEFINITION us90b10.y1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
 IMAGE:3325627 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
 ACCESSION BE334877
 VERSION BE334877.1 GI:9208653
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 585)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 This clone is available royalty-free through LILNL; contact the
 IMAGE Consortium (info@image.lilnl.gov) for further information.
 MGI:1069791
 Seq primer: -40RP from Gibco
 High quality sequence stop: 465.

FEATURES

Source

1..585
 Location/Qualifiers
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_image="3325627"
 /clone_lib="Soares_mammary_gland_NMLMG"
 /sex="female (lactating)"
 /tissue_type="mammary gland"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from mammary
 gland tissue from a lactating female, and was then primed
 with a Not I - oligo(dT) primer. Double-stranded cDNA was
 ligated to Eco RI adaptors (Pharmacia), digested with Not
 I and cloned into the Not I and Eco RI sites of the
 modified pT7T3 vector. Library is normalized. Library
 was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 162 a 157 c 172 g 93 t 1 others
 ORIGIN
 alignment_scores:
 Quality: 688.00 Length: 124
 Ratio: 5.548 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 98.387
 alignment_block:
 US-09-327-750D-12 x BE334877 ..
 Align seg 1/1 to: BE334877 From: 1 to: 585

1 MetAlaAsnValHisGlnGluAsnGluGluMetGluGlnProLeuGlnAs 17
 166 ATGGCCAATGTCCACCAGAAACGAGAGATGGAGCAGCCCTGCAGAA 215
 17 nGlyGluGluAspArgProValGlyGlyGlyGlyHisGlnProAlaG 34
 216 TGGACAGGAAGACCGCCCTGTGGAGGAGGTGAGGGCCACCACTGCTG 265
 34 lYAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArg 50
 266 CAAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 315
 51 GlyGlnAlaArgLeuAlaProAsnPheArgTTPAlaIleProAsnAr 67
 316 GCCAGGCTCGCGGACTTCCCTTAACTTCCGATGGCCATTCCCAACAG 365
 67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
 366 GCAGATGAATGACGGTGTGGGTGGAGATGGAGATGATGGAATGTTCA 415
 84 etGluGluMetArgGluIleArgArgLysLeuArgGluLeuGlnLeuArg 100
 416 TGGAGGAGATGAGAGATCCGAGAAAGCTTAGGGAGCTACAGCTGAGA 465
 101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisAsnHisH 117
 466 AATTGCTTACGATCTTATGGGGAGCTGTCTAACCAACCAAGATCACCA 515
 117 sAspGluPheCysLeuMetPro 124
 516 TGATGAATTCGCTTATGCGCT 537

seq_name: gb_est2:w64711

seq_documentation_block:
 LOCUS W64711 590 bp mRNA EST 10-JUN-1996
 DEFINITION md71q01.r1 Soares mouse embryo NbME13 5 14.5 Mus musculus cDNA
 clone IMAGE:373872 5' similar to gb:M38188 OVARIAN GRANULOSA CELL
 13.0 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.

ACCESSION W64711
 VERSION W64711.1 GI:1372353
 KEYWORDS EST.
 SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 590)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE

The WashU-HHMI Mouse EST Project

JOURNAL

Unpublished (1996)

COMMENT

Contact: Marra M/Mouse EST Project
 WashU-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800

Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:235304

Seq primer: EFPrimer
 High quality sequence stop: 348.

FEATURES

Source
 Location/Qualifiers
 1..590
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:373872"
 /clone_lib="Soares mouse embryo NbME13.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /note="Vector: p7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TCTTACCAATCTGAGTGGAGCGCGCGGAAATTTTCTTTTCTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
 State Univ., from 2]; double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia). digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT73 vector. Library went through one round of
 normalization. " was constructed by Bento Soares and
 M.Fatima Bonaldo."

BASE COUNT 165 a 155 c 166 g 104 t
 ORIGIN

alignment_scores:

Quality: 688.00 Length: 124
 Ratio: 5.548 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 98.387

alignment_block:

US-09-327-750D-12 x W64711 ..
 Align seq 1/1 to: W64711 from: 1 to: 590

```

1 MetAlaAsnValHisGlnGluAsnGluMetClnProLeuGlnAs 17
134 ATGCCCAATCTCACCAGGAAACCAAGAGATGGAGCCCTGCAGAA 183
17 nGlyGluGluAspArgProValGlyGlyGluGlyHisGlnProAlaG 34
184 TGGACAGGAAGACCGCCTGTGGGAGGAGGTGAGGGCCACCGCTGCTG 233
34 lyAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisHisArgArg 50
234 CAACAACAACAACAACAACAACCAACCAACCAACCAACCAACCA 283
51 GlyClnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 67
284 GGCAGGCTCGCGACTTGCCTTAACTCCGATGGGCCATCCCAACAG 333
67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
334 GCAGATGAATGACGGTGTGGTGAGATGGAGATGATATGGAATGTCA 383
84 etGluGluMetArgGluIleArgArgLysLeuArgGluLeuGlnLeuArg 100
384 TGGAGGAGATGAGAGAGATCCGGAGAAAGCTTAGGGAGCTACAGCTCAGA 433
101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisH 117
434 AATGTCTACGCATCCCTTATGGGGAGCTGTCTAAACCACCACGATCACCA 483
117 sAspGluPheCysLeuMetPro 124
```

|||||
 484 TGATGAATCTGCCTTATGCCT 505

seq_name: gb_est2:BG088461

seq_documentation_block:

LOCUS BG088461 599 bp mRNA EST 26-JAN-2001
 DEFINITION H3153D07-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
 H3153D07 5', mRNA sequence.
 ACCESSION BG088461
 VERSION BG088461.1 GI:12571025
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 599)
 AUTHORS Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
 ,T.S., Carter,M.G. and Ko,M.S.H.
 TITLE Verification and initial annotation of NIA mouse 15K cDNA clone set
 JOURNAL Unpublished (2001)
 COMMENT Other_ESTs: H3153D07-3
 Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please
 visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
 Plate: H3153 row: D column: 07
 Seq primer: -21M13 Reverse
 High quality sequence stop: 599
 POLYA-No.

FEATURES

Location/Qualifiers

1..599
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="H3153D07"
 /clone_lib="NIA Mouse 15K cDNA Clone Set"
 /sex="Clones arrayed from a variety of cDNA libraries"
 /dev_stage="Clones arrayed from a variety of cDNA
 libraries"
 /lab_host="DH10B"
 /note="Vector: pSPORT1; Site_1: Salt; Site_2: NotI; This
 clone is among a rearranged set of 15,247 clones from 11
 embryo cDNA libraries (including preimplantation stage
 embryos from unfertilized egg to blastocyst, embryonic
 part of E7.5 embryos, extraembryonic part of E7.5 embryos
 , and E12.5 female mesonephros/gonad) and one newborn
 ovary cDNA library. Average insert size 1.5 kb. All
 source libraries are cloned unidirectionally with Oligo(dT
)-Not primers. References include: (1) Genome-wide
 expression profiling of mid-gestation placenta and embryo
 using a 15,000 mouse developmental cDNA microarray, 2000,
 Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)
 Large-scale cDNA analysis reveals phased gene expression
 patterns during preimplantation mouse development, 2000,
 Development, 127: 1737-1749; (3) Genome-wide mapping of
 unselected transcripts from extraembryonic tissue of
 7.5-day mouse embryos reveals enrichment in the t-complex
 and under-representation on the X chromosome, 1998, Hum
 Mol Genet 7: 1967-1978."

BASE COUNT 163 a 163 c 176 g 97 t
 ORIGIN

alignment_scores:

Quality: 688.00 Length: 124
 Ratio: 5.548 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 98.387

alignment_block:

US-09-327-750D-12 x BG088461

Align seg 1/1 to: BG088461 from: 1 to: 599

```

1  MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
|||||
175 ATGGCCCAATGTCACACGAGAAACGAGAGATGGAGCAGCCCTGCAGAA 224
|||||
17  nGlyGluGluAspArgProValGlyGlyGlyGlyGlyGlnProAlaAg 34
|||||
225 TGGACAGGAAGACCCCTGTGGAGGAGGTGAGGCCACACGACCTGCTG 274
|||||
34  lYAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisHisArg 50
|||||
275 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 324
|||||
51  GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaLeuProAsnAr 67
|||||
325 GGCCAGGCTCGCGGACTTGCCTTAACCTTCGATGGGCCATTCCTCAACAG 374
|||||
67  gGlnMetAsnAspGlyLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 84
|||||
375 GCAGATGAATGACGGCTTGGGTGGAGATGGAGATGATGAAATGTTCA 424
|||||
84  etGluGluMetArgGluLeuArgArgLysLeuArgGluLeuGlnLeuArg 100
|||||
425 TGGAGGAGATGAGAGAGATCCGGAGAAAGCTTAGGGAGCTACAGCTGAGA 474
|||||
101 AsnCysLeuArgIleLeuMetGlyGlyLeuSerAsnHisHisAspHisHi 117
|||||
475 AATTGCTACGCATCCTTATGGGGAGCTGTCTAACCCACACGATCACCA 524
|||||
117 sAspGluPheCysLeuMetPro 124
|||||
525 TCATGAATTCGCTTATGCGCT 546
|||||

```

seq_name: gb_est1:AA272375

seq_documentation_block: 26-MAR-1997
LOCUS AA272375 612 bp mRNA EST
DEFINITION vb62407.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
clone IMAGE:761628 5' similar to gb:M38188 OVARIAN GRANULOSA CELL
13.0 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.

ACCESSION AA272375
VERSION AA272375.1 GI:1910706
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 612)
Marta, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Willson, R. and
Waterston, R.

TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu

This clone is available royalty-free through LINL : contact the
IMAGE Consortium (info@image.lnli.gov) for further information.
MGI:462548
Seq primer: -28ml3 rev2 Et from Amersham
High quality sequence stop: 507.
Location/Qualifiers

FEATURES
source
1..612
/organism="Mus musculus"

/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:761628"
/clone_lib="Barstead mouse pooled organs MPLRB4"
/sex="mixed"
/tissue_type="pooled organs"
/dev_stage="7 day"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACGAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[GTGGATTCGGTACC], digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead."

BASE COUNT 168 a 167 c 175 g 102 t
ORIGIN

alignment_scores:
Quality: 688.00 Length: 124
Ratio: 5.548 Gaps: 0
Percent Similarity: 100.000 Percent identity: 98.387
alignment_block:
US-09-327-750D-12 x AA272375 ..

Align seg 1/1 to: AA272375 from: 1 to: 612

```

1  MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
|||||
176 ATGGCCAATGTCACACGAGAAACGAGATGGAGCAGCCCTGCAGAA 225
|||||
17  nGlyGluGluAspArgProValGlyGlyGlyGlyGlyGlnProAlaAg 34
|||||
226 TGGACAGGAAGACCCCTGTGGAGGAGGTGAGGCCACACGACCTGCTG 275
|||||
34  lYAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisHisArg 50
|||||
276 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 325
|||||
51  GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaLeuProAsnAr 67
|||||
326 GGCCAGGCTCGCGGACTTGCCTTAACCTTCGATGGGCCATTCCTCAACAG 375
|||||
67  gGlnMetAsnAspGlyLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 84
|||||
376 GCAGATGAATGACGGGTGGGTGGAGATGGAGATGATGAAATGTTCA 425
|||||
84  etGluGluMetArgGluLeuArgArgLysLeuArgGluLeuGlnLeuArg 100
|||||
426 TGGAGGAGATGAGAGAGATCCGGAGAAAGCTTAGGGAGCTACAGCTGAGA 475
|||||
101 AsnCysLeuArgIleLeuMetGlyGlyLeuSerAsnHisHisAspHisHi 117
|||||
476 AATTGCTACGCATCCTTATGGGGAGCTGTCTAACCCACACGATCACCA 525
|||||
117 sAspGluPheCysLeuMetPro 124
|||||
526 TGATGAATTCGCTTATGCGCT 547
|||||

```

seq_name: gb_est2:W81757

seq_documentation_block:
LOCUS W81757 616 bp mRNA EST
DEFINITION me95d06.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
clone IMAGE:403307 5' similar to gb:M38188 OVARIAN GRANULOSA CELL
13.0 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION W81757
VERSION W81757.1 GI:1392776
KEYWORDS EST.
SOURCE house mouse.

BASE COUNT 171 a 162 c 181 g 112 t 2 others
ORIGIN

alignment_scores:

Quality: 688.00 Length: 124
Ratio: 5.548 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.387

alignment_block:

US-09-327-750d-12 x AW476468 ..

Align seg 1/1 to: AW476468 from: 1 to: 628

1 MetaAAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17

138 ATGGCCAAATGTCCACCAGGAAACCAAGAGATGGAGCCCTCGAGAA 187

17 nGlyGluGluAspArgProValGlyGlyGlyGluHisGlnProAlaG 34

188 TGCAGCAGGAAGACCCGCTGTGGGAGAGGTGAGGCCACCAACGCTGCTG 237

34 lYAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArg 50

238 CAAC 287

51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 67

288 GCCCAGGCTCGCGACTTGGCCCTAACCTCCGATGGCCATCCCAACAG 337

67 gClnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84

338 CCAGATGAATGACGGGTGGGTGGAGATGGAGTGGATGGAATGTTC 387

84 eGluGluMetArgGluIleArgArgLysLeuArgGluLeuGlnLeuArg 100

398 TGGACGAGATGACAGAGATCCGAGAAAGCTTAGGGAGCTACAGCTGAGA 437

101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisH 117

438 AATTGCTACGCATCTTATGGGGAGCTGTCTACACACACGATCACCA 487

117 sAspGluPheCysLeuMetPro 124

488 TGATGAATTCGCTTATGCCT 509

seq_name: gb_hlc-AK004531

seq_documentation_block:

LOCUS AK004531 809 bp mRNA HTC 05-JUL-2001
DEFINITION Mus musculus 18 days embryo cDNA, RIKEN full-length enriched library, clone:1190011J23, full insert sequence.

ACCESSION AK004531

VERSION AK004531.1 GI:12835760

KEYWORDS CAP trapper.

SOURCE Mus musculus (strain:C57BL/6J) 18 days embryo cDNA to mRNA, clone:1190011J23.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 809)

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Methods in enzymology. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2 (bases 1 to 809)

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome research. 10 (10), 1617-1630 (2000)

MEDLINE
PUBMED
REFERENCE
AUTHORS

20499374
11042159

3 (bases 1 to 809)

Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome research. 10 (11), 1757-1771 (2000)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

20530913
11076861

4 (bases 1 to 809)

The RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

5 (bases 1 to 809)

Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
Arakawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Horl,F.,
Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,
Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T.,
Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,
Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A.,
Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission

TITLE
JOURNAL

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT

Please visit our web site (http://genome.gsc.riken.go.jp/) for

further details.

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
15' GAGAGAGAGCGCCGCACTCGAGTTTCTTTTCTTTTNN 3', cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 5.0. Second strand cDNA
was prepared with the primer adapter of sequence[5'
GAGAGAGAGAGGATCCAGAGCTCAATTAATTAATTAACCCCCCCC 3']. cDNA was
cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end:

XhoI. Host: SOLR.

Location/Qualifiers

1. .809
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGD:MGI:1338016"
/db_xref="MGD:MGI:1892692"
/clone="1190011J23"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="18 days embryo"
181. .624
/note="putative"
/codon_start=1
/protein_id="BAB23350.1"
/db_xref="GI:12835761"
/translation="MANVHQENDEMEQPLQNGQEDRPVGGGEGHQPAANNNNNNHNNH
HNHRRGQARRLAPNRWAIPIRQMDNGDGGDDMMEMEMEMERLRRLRLQLRNC
LRLMLGELSNHHDDHDEFCLMPWLSRFPPEIHTVTTPAVALFLAFS"

FEATURES
source

CDS

polyA_signal 790..795
 /note="putative"
 polyA_site 809
 /note="putative"
 BASE COUNT 208 a 211 c 227 g 163 t
 ORIGIN

alignment_scores:
 Quality: 688.00 Length: 124
 Ratio: 5.548 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 98.387

alignment_block:

US-09-327-750D-12 x AK004531 ..

Align seg 1/1 to: AK004531 from: 1 to: 809

1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
 |||||
 181 ATGGCCAAATGTCACAGGAAGAAAGAGATGGAGCAGCCCTGCAGAA 230
 |||||
 17 nGlyGluGluAspArgProValGlyGlyGlyGluGlyHisGlnProAlaG 34
 |||||
 231 TGGACAGGAAGACCGCCCTGTGGGAGGAGGTGAGGGCCACAGCCTGCTG 280
 |||||
 34 lYAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArg 50
 |||||
 281 CAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 330
 |||||
 51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 67
 |||||
 331 GGCCAGGCTCGCCGACCTGCCCTTAACCTCGGATGGGCCATTCACCAACAG 380
 |||||
 67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
 |||||
 381 GCAGATCAATGACGGTGGGTGGGAGATGGAGATGATGAAATGTTCAC 430
 |||||
 84 etGluGluMetArgGluIleArgArgLysLeuArgGluLeuGluLeuArg 100
 |||||
 431 TGGAGGAGATGAGAGAGATCCGGAGAAGCTTAGGGAGCTACAGCTGAGA 480
 |||||
 101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisHi 117
 |||||
 481 AATTGTCTACGCATCCTTATGGGGAGCTGTCTAACCCACCCAGATCACCA 530
 |||||
 117 sAspGluPheCysLeuMetPro 124
 |||||
 531 TGATGAATTCTGCCTTATGCCT 552
 |||||

seq_name: gb_est1:AA253897

seq_documentation_block: 468 bp mRNA EST 13-MAR-1997
 LOCUS AA253897
 DEFINITION my31e10.r1 Barstead mouse pooled organs MPLRB4 Mus musculus cDNA
 clone IMAGE:697482 5', similar to gb:M38188 OVARIAN GRANULOSA CELL
 13.0 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.

ACCESSION AA253897

VERSION AA253897.1 GI:1888253

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota: Metazoa: Chordata: Craniata; Vertebrata: Euteleostomi;
 Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS 1 (bases 1 to 468)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.

TITLE The WashU-HHMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.

MGI:431042

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 455.

Location/Qualifiers

1..468

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone_lib="Barstead mouse pooled organs MPLRB4"

/sex="mixed"

/tissue_type="pooled organs"

/dev_stage="7 day"

/lab_host="DH10B"

/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
 a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5',
 TGTTACGAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 [GTTGATTCGGTACC], digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pT73 vector.
 Library constructed by Bob Barstead."

BASE COUNT 140 a 126 c 120 g 82 t

ORIGIN

alignment_scores:

Quality: 687.00 Length: 124

Ratio: 5.540 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 97.581

alignment_block:

US-09-327-750D-12 x AA253897 ..

Align seg 1/1 to: AA253897 from: 1 to: 468

1 MetAlaAsnValHisGlnGluAsnGluMetGluGlnProLeuGlnAs 17
 :|||||
 49 ATGCCCAATGTCCACCAGGAACAGAGAGATGGAGCAGCCCTGCAGAA 98
 :|||||
 17 nGlyGluGluAspArgProValGlyGlyGlyGlyHisGlnProAlaG 34
 :|||||
 99 TGGACAGGAAGACCCCTCTGGGAGGAGGTGAGGGCCACAGCCTGCTG 148
 :|||||
 34 lYAsnAsnAsnAsnAsnAsnHisAsnHisAsnHisAsnHisAsnHisArg 50
 :|||||
 149 CAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 198
 :|||||
 51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 67
 :|||||
 199 GGCCAGGCTCGCCGACCTTGGCCCTTAACCTCGGATGGGCCATTCACCAACAG 248
 :|||||
 67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
 :|||||
 249 GCAGATCAATGACGGTGGGTGGAGATGGAGATGATGATGAAATGTTCAC 298
 :|||||
 84 etGluGluMetArgGluIleArgArgLysLeuArgGluGluLeuLeuArg 100
 :|||||
 299 TGGAGGAGATGAGAGAGATCCGGAGAAGCTTAGGGAGCTACAGCTGAGA 348
 :|||||
 101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisHi 117
 :|||||
 349 AATTGTCTACGCCTTATGGGGAGCTGTCTAACCCACCCAGATCACCA 398
 :|||||
 117 sAspGluPheCysLeuMetPro 124

399 TGATGAATCTGCCTTATGCCT 420

seq_name: qb htc:AK010500

seq_documentation_block:
LOCUS AK010500 845 bp mRNA 05-JUL-2001
DEFINITION Mus musculus ES cells cDNA, RIKEN full-length enriched library,
clone:2410015K23, full insert sequence.
ACCESSION AK010500
VERSION AK010500.1 GI:12845991
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) ES cells cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library
clone:2410015K23.

ORGANISM	Mus musculus
REFERENCE	1. (bases 1 to 845)
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Methods in enzymology. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636

FEATURES

101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisH117
 478 AATTGTCTACGCATCTTATGGGAGCTGTCTAAACCACCGATCACCA 527
 117 saspGluPheCysLeuMetPro 124
 528 TGATGAATTCTGCCTTATGCCT 549

seq_name: gb_est1:A1152323

seq_documentation_block: 430 bp mRNA EST 30-SEP-1998
 LOCUS A1152323
 DEFINITION u079a02.r1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
 IMAGE:1477034 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.

ACCESSION A1152323

VERSION A1152323.1 GI:3680792

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE
 AUTHORS
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 430)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

TITLE The WashU-HMI Mouse EST Project

JOURNAL

COMMENT

Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:925390

Seq primer: -28m13 rev2 ET from Amersham

High quality sequence stop: 419.

FEATURES

source

Location/Qualifiers
 1..430
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:1477034"
 /clone_lib="Soares_mammary_gland_NMLMG"
 /sex="female (lactating)"
 /tissue_type="mammary gland"
 /lab_host="DH10B"

/note="vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from mammary
 gland tissue from a lactating female, and was then primed
 with a Not I - oligo(dT) primer. Double-stranded cDNA was
 ligated to Eco RI adaptors (Pharmacia), digested with Not
 I and cloned into the Not I and Eco RI sites of the
 modified pT7T3 vector. Library is normalized. Library
 was constructed by Bento Soares and M. Fatima Donaldo."

BASE COUNT 135 a 111 c 113 g 71 t

ORIGIN

alignment_scores:

Quality: 681.00 Length: 124
 Ratio: 5.537 Gaps: 0
 Percent Similarity: 99.194 Percent Identity: 97.581

alignment_block:

US-09-327-750D-12 x A1152323

Align seg 1/1 to: A1152323 from: 1 to: 430

1 MetAlaAsnValHisGlnGlnAsnGluGluMetGluGlnProLeuGlnAs 17
 54 ATGGCCAATGTCCACAGGAAACGAAGAGATGGAGCAGCCCTGCAGAA 103
 17 nGlyGluGluAspArgProValGlyGlyGlyGluGlnProAlaG 34
 104 TGGACAGGAAGACCGCCCTGTGGATGGAGGTGAGGCCACCGCCTGCTG 153
 34 lYAsnAsnAsnAsnAsnAsnHisHisHisHisHisHisHisHisHis 50
 154 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 203
 51 GlyGlnAlaArgArgLeuAlaProAsnPheArgTTPAlaIleProAsnAr 67
 204 GCCCAGGCTGCCCGACTTCCCTTAACCTCCGATGGCCCATTTCCCAACAG 253
 67 gGlnMetAsnAspGlyLeuGlyGlyAspGlyAspMetGluMetPheM 84
 254 GCAGATGAATGACGGGTGGGTGGAGATGGAGATGATATGGAATGTTC 303
 84 etGluGluMetArgGluLeuLeuArgLysLeuArgGluLeuGlnLeuArg 100
 304 TGGAGGAGATGAGAGAGATCCCGAGAAAGCTTAGGGAGCTACAGCTGAGA 353
 101 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisH 117
 354 AATTGTCTACGCATCTTATGGGAGCTGTCTAACCCACCGATCACCA 403
 117 saspGluPheCysLeuMetPro 124
 404 TGATGAATTCTGCCTTATGCCT 425

seq_name: gb_est1:AW908751

seq_documentation_block:

LOCUS AW908751 503 bp mRNA EST 25-MAY-2000
 DEFINITION uf57a05.y1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
 IMAGE:1515440 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.

ACCESSION AW908751

VERSION AW908751.1 GI:8073984

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 503)

REFERENCE NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

TITLE Unpublished (1997)

JOURNAL Other ESTs: uf57a05.x1

COMMENT Contact: Robert Strausberg, Ph.D.

Email: ccapbs-r@mail.nih.gov

This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:942292

Seq primer: -40RP from Gibco

High quality sequence stop: 468.

FEATURES

source

Location/Qualifiers
 1..503
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:1515440"
 /clone_lib="Soares_mammary_gland_NMLMG"
 /sex="female (lactating)"
 /tissue_type="mammary gland"
 /lab_host="DH10B"

/note="vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from mammary
 gland tissue from a lactating female, and was then primed
 with a Not I - oligo(dT) primer. Double-stranded cDNA was
 ligated to Eco RI adaptors (Pharmacia), digested with Not

I and cloned into the Not I and Eco RI sites of the
modified p7T3 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo.
BASE COUNT 132 a 135 c 130 g 105 t 1 others
ORIGIN

```
alignment_scores:
  Quality: 679.00      Length: 122
  Ratio: 5.566        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 98.361

alignment_block:
US-09-327-750D-12 x AW908751 ..
Align seg 1/1 to: AW908751 from: 1 to: 503

3 AsnValHisGlnGluAsnGluGluMetGluGlnProLeuGlnAsnGlyG1 19
|||||
3 AATGTCCACCAGGAACGAAGAGATGGAGCAGCCCTGCAGATGGACA 52
|||||
19 uGluAspArgProValGlyGlyGluGlyHisGlnProAlaGlyAsnA 36
:|||||
53 GGAAGACCGCCCTGTGGGAGGAGGTGAGGGCCACCAGCCTGCTCCAACA 102
|||||
36 snAsnAsnAsnHisAsnHisAsnHisAsnHisArgArgGlyGln 52
|||||
103 ACAACAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAAGGCCAG 152
|||||
53 AlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnArgGlnMe 69
|||||
153 GCTCGCGGACTTGCCCTAACTTCGGATGGGCCATTCCCAACAGGCAGAT 202
|||||
69 tAsnAspGlyLeuGlyClyAspGlyAspMetGluMetPheMetGluG 86
|||||
203 GAATGACGGGTTGGGTGGAGATGGAGATGATGAAATGTTTCATGGAGG 252
|||||
86 luMetArgGluIleArgArgGlyLeuArgGluLeuGlnLeuArgAsnCys 102
|||||
253 AGATGAGAGAGATCCGGAGAAAGCTTAGGGAGCTACAGCTGAGAAATGT 302
|||||
103 LeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisAspG1 119
|||||
303 CTACGCATCCTTATGGGGAGGCTGTCTAACCAACCAACCAACCAACCAATGTA 352
|||||
119 uPheCysLeuMetPro 124
|||||
353 ATTCTGCCTTATGCCT 368
```

OM of: US-09-327-750D-13 to: GenEmbl:* out_format : pfs
Date: Mar 11, 2002 3:33 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=Cgnt2_1/USPTO.spool/US09327750/runat_11032002_101154_20319/app_query.fasta_1.1472
-DB=GenEmbl -OFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTPWT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09327750 -CGNT1_1_8673
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-13
Query length: 111
Database: GenEmbl:*
Database sequences: 1472140
Database length: 34134837
Search time (sec): 4557.230000

score_list:

Sequence	Strd	Orig	zScore	Escore	Len	Documentation
gb_pr:BC003190	+	609.00	947.76	1.6e-44	793	BC003190 Homo sapiens, p75NTR-as
gb_pr:HUMOGC	+	609.00	946.94	1.8e-44	891	M38188 Human unknown protein f
gb_pr:AF187066	+	609.00	946.94	1.8e-44	891	AF187064 Homo sapiens p75NTR-as
gb_pr:AF187066	+	567.50	884.08	5.7e-41	700	AF187066 Mus musculus p75NTR-as
gb_ro:AF097440	+	567.50	882.68	6.9e-41	854	AF097440 Mus musculus brain exp
gb_ro:AF187065	+	538.50	841.06	1.4e-38	519	AF187065 Rattus norvegicus p75N
gb_sts:G72708	-	418.00	655.08	3.3e-28	421	G72708 MARCH 4953-4954:991939031
gb_pr:HSV23108	+	300.00	438.72	3.7e-16	45678	G72708 MARCH 4953-4954:991939031
gb_pr:AK000959	+	277.00	428.26	1.4e-15	1229	AK000959 Homo sapiens cDNA FLJ
gb_pr:AX100231	+	277.00	427.53	1.5e-15	1364	AX100231 Sequence 32 from Paten
gb_pr:HS635619	+	277.00	399.99	5.3e-14	69648	AL035494 Human DNA sequence f
gb_sts:G35294	-	249.50	392.10	1.5e-13	477	G35294 human SPS SHGC-37409, se
gb_pr:HS17487	+	241.50	342.36	8.6e-11	98274	G29755 Human DNA sequence fro
gb_pr:AF183416	+	194.50	304.47	1.1e-08	642	AF183416 Homo sapiens ovarian g
gb_pr:AF220189	+	194.50	302.69	1.4e-08	838	AF220189 Homo sapiens uncharact
gb_pr:HS19894	+	175.00	276.34	4.1e-07	35714	AL008708 Human DNA sequence f
gb_pr:AF237783	+	175.00	292.12	5.4e-08	791	AF237783 Homo sapiens brain-exp
gb_pr:AF097439	+	175.00	273.51	5.9e-07	785	AF097439 Mus musculus brain exp
gb_pr:AF251053	+	175.00	273.11	6.2e-07	744	AF251053 Homo sapiens X-linked
gb_pr:AX078272	+	175.00	271.79	7.3e-07	898	AX078272 Sequence 76 from Paten
gb_pr:HSV87088	+	175.00	246.92	1.8e-05	31321	G72708 MARCH 4953-4954:991939031
gb_pr:AL133348	-	175.00	245.11	2.2e-05	40584	AL133348 Human DNA sequence f
gb_hgt:HSU8081	+	175.00	245.03	2.3e-05	41029	AL022169 Homo sapiens chromos
gb_ro:AF097438	+	158.50	246.63	1.8e-05	835	AF097438 Mus musculus brain exp
gb_ro:AF051347	+	158.50	246.61	1.8e-05	838	AF051347 Mus musculus REX-3 mRN
gb_ro:AF097437	+	158.50	239.63	4.5e-05	2269	AF097437 Mus musculus Bex1 pro
gb_pr:HSU10564	+	156.00	215.00	0.0011	43952	G72708 MARCH 4953-4954:991939031
gb_hgt:AL357046	+	140.50	182.07	0.0728	154577	AL357046 Homo sapiens chromo
gb_pr:AC009282	+	140.50	182.07	0.0728	154595	AC009282 Homo sapiens clone
gb_pr:AL35305	+	140.50	180.82	0.0854	185257	AL35305 Human DNA sequence
gb_sts:G24641	-	120.50	191.06	0.0230	504	G24641 human SPS WI-11354, sequ
gb_pr:BC005988	+	105.00	161.09	1.07	1163	BC005988 Homo sapiens, clone 1
gb_pr:HSU10564	+	105.00	135.67	27.95	43952	G72708 MARCH 4953-4954:991939031
gb_ro:BC003254	+	97.00	142.25	12.01	2901	BC003254 Mus musculus, Similar
gb_ba:MXAC000101	+	95.00	138.36	19.78	3622	AC000101 00107, complete sequ
gb_hgt:AC012237	+	95.00	112.55	541.57	129339	AC012237 Homo sapiens clone
gb_hgt:AL592044	+	95.00	108.17	950.81	242092	AL592044 Homo sapiens chromo
gb_pr:AK026349	+	91.50	139.83	16.39	1208	AK026349 Homo sapiens cDNA: FI
gb_pr:HSU177E8	+	91.50	115.18	386.58	40822	G72708 MARCH 4953-4954:991939031
gb_pr:HSU823F3	-	91.50	109.53	797.75	91480	AL079333 Human DNA sequence f

gb_hgt:LMFLCHR16_01 - 91.00 107.47 1.0e+03 110000 ! Continuation (2 of 11) of
gb_ba:AVIALGEB + 88.50 117.18 299.25 15759 ! L39096 Azotobacter vinelina
gb_pl:TIN251864 + 87.50 127.24 82.39 3000 ! AJ251864 Tolypocladium infl
gb_ba:AE002479 - 87.50 118.39 256.15 10612 ! AE002479 Neisseria meningi
gb_ba:NMA52491 - 87.50 94.33 5.6e+03 329861 ! AL162756 Neisseria mening
seq_name: gb_pr:BC003190
seq_documentation_block:
LOCUS BC003190 793 bp mRNA PRI 12-JUL-2001
DEFINITION Homo sapiens, p75NTR-associated cell death executor; ovarian
granulosa cell protein (13kD), clone MGC:802 IMAGE:3357965, mRNA,
complete cds.
ACCESSION BC003190
VERSION BC003190.1 GI:13112030
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 793)
AUTHORS Strausberg, R
TITLE Direct Submission
JOURNAL Submitted (13-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 6 Row: k Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7657043.

FEATURES

source

1..793
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="LocusID:27018"
/db_xref="taxon:9606"
/clone="MGC:802 IMAGE:3357965"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH-MGC_16"
/lab_host="DH10B-R"
/note="vector: pOTB7"
196..531
/codon_start=1
/product="p75NTR-associated cell death executor; ovarian
granulosa cell protein (13kD)"
/protein_id="AAH03190.1"
/db_xref="GI:13112031"
/translation="MANTHOENMEQPMQNGEDRPLGGEGHQPNRRGQARRLA
PNFRWAPNQRINDMGSGDDMEIFMEERKLRELQLRNLCLRLMSELSNHHD
HHDEFCLMP"

CDS

225 a 170 c 205 g 193 t

BASE COUNT
ORIGIN

```
alignment_scores:
  Quality: 609.00      Length: 111
  Ratio: 5.486         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-327-750D-13 x BC003190 ..
  Align seg 1/1 to: BC003190 from: 1 to: 793

1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
196 ATGCCAAATATTCACAGGAAACGAGAGATGGAGCGCTATGCAGAA 245
17 nGlyGluAspArgProLeuGlyGlyGlyGluGlyHisGlnProAlaG 34
246 TGGAGAGAGACCGCCCTTTGGGAGGAGGTGAAGCCACCGCTGCAG 295
34 lyAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAla 50
296 GAAATCGACGGGACAGCTCGCGACTTGCCCTAATTTTCGATGGGCC 345
51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspMe 67
346 ATACCAATAGGCAGATCAATGATGGGATGGTGGAGATGGAGATGAT 395
67 tGluIlePheMetGluMetArgGluIleArgGlyLysLeuArgGluL 84
396 GGAATATTCATGGAGGAGATGAGAGAAATCAGAAGAAACTTAGGGAGC 445
84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
446 TGCAGTTGAGGAATGTCGCGTATCTTATGGGGAGCTCTCTAATCAC 495
101 HisAspHisHisAspGluPheCysLeuMetPro 111
496 CATGACCATCATGAATTTTGCCTTATGCCT 528

seq_name: gb_pr:HUOGC

seq_documentation_block:
  LOCUS HUOGC 891 bp mRNA 07-MAR-1995
  DEFINITION Human unknown protein from clone pHR74 mRNA, complete cds.
  ACCESSION M38188 X56942
  VERSION M38188.1 GI:189378
  KEYWORDS
  SOURCE Human ovarian granulosa cell line, cDNA to mRNA.
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    Rapp.G., Freudenstein,J., Klaudiny,J., Mucha,J., Wempe,F.,
    Zimmer,M. and Scheit,K.H.
  TITLE Characterization of three abundant mRNAs from human ovarian
    granulosa cells
  JOURNAL DNA Cell Biol. 9 (7), 479-485 (1990)
  MEDLINE 91025550
  COMMENT Draft entry and computer-readable sequence for [DNA 9, 479-485
    (1990)] kindly submitted
    by K.H.Scheit, 27-AUG-1990.
  FEATURES
    Location/Qualifiers
      1..891
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="pHR74"
        /cell_type="granulosa"
        /tissue_type="ovary"
        <1..891
        /note="protein of unknown function"
      312..647
        /note="protein of unknown function"
        /codon_start=1
        /protein_id="AAA63232.1"

CDS
  /db_xref="GI:189379"
  /translation="MANIHQENEEQPNQCEEDRPLGGEGGHQHPAGNRRQCARRLA
  PNFRWAIPINRQINDMGGGDDMEIFMEEMREIRKRLQLRNLCLILMUGELSNHHD
  HHDEFCLMP"
  361..534
  /note="protein of unknown function"
  /codon_start=1
  /protein_id="AAA63233.1"
  /db_xref="GI:189380"
  /translation="MERKTALEWEEVKATSLQEIQDGLADLPLIFDGPYPPIGRSMWG
  VEMEMIKWYSRR"

BASE COUNT 251 a 182 c 224 g 234 t
ORIGIN

alignment_scores:
  Quality: 609.00      Length: 111
  Ratio: 5.486         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-327-750D-13 x HUOGC ..
  Align seg 1/1 to: HUOGC from: 1 to: 891

1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
312 ATGCCAAATATTCACAGGAAACGAGAGATGGAGCGCTATGCAGAA 361
17 nGlyGluAspArgProLeuGlyGlyGlyGluGlyHisGlnProAlaG 34
362 TGGAGAGAGACCGCCCTTTGGGAGGAGGTGAAGCCACCGCTGCAG 411
34 lyAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAla 50
412 GAAATCGACGGGACAGGCTCGCGACTTGCCCTAATTTTCGATGGGCC 461
51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspMe 67
462 ATACCAATAGGCAGATCAATGATGGGATGGTGGAGATGGAGATGATAT 511
67 tGluIlePheMetGluMetArgGluIleArgGlyLysLeuArgGluL 84
512 GGAATATTCATGGAGGAGATGAGAGAAATCAGAAGAAACTTAGGGAGC 561
84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
562 TGCAGTTGAGGAATGTCGCGTATCTTATGGGGAGCTCTCTAATCAC 611
101 HisAspHisHisAspGluPheCysLeuMetPro 111
612 CATGACCATCATGAATTTTGCCTTATGCCT 644

seq_name: gb_pr:AF187064

seq_documentation_block:
  LOCUS AF187064 891 bp mRNA 11-JUN-2000
  DEFINITION Homo sapiens p75NTR-associated cell death executor (NADE) mRNA,
    complete cds.
  ACCESSION AF187064
  VERSION AF187064.1 GI:8452893
  KEYWORDS
  SOURCE human.
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 (bases 1 to 891)
    Mukai,J., Hachiya,T., Shoji-Hoshino,S., Kimura,M.T., Nadano,D.,
    Suvanto,P., Hanaoka,T., Li,Y., Irie,S., Greene,L.A. and Sato,T.A.
    NADE, a p75NTR-associated cell death executor, is involved in
    signal transduction mediated by the common neurotrophin receptor
    p75NTR
  JOURNAL J. Biol. Chem. 275 (23), 17566-17570 (2000)
```



```
|||||
427 TGGAGGAGATGAGAGATCGGAGAAAGCTTAGGGAGCTACAGCTGAGA 476
|||||
88 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisHi 104
|||||
477 AATTGCTACGATCCTTTATGGGGAGCTGTCTAACCCACCACCATCA 526
|||||
```

```
104 sAspGluPheCysLeuMetPro 111
|||||
527 TGATGAATTCGCTTATGCCT 548
|||||
```

seq_name: gb_ro:AF097440

```
seq_documentation_block:
LOCUS AF097440 854 bp mRNA 13-APR-1999
DEFINITION Mus musculus brain expressed X-linked protein 3 (Bex3) mRNA,
complete cds.
ACCESSION AF097440
VERSION AF097440.1 GI:4580593
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
Hum. Mol. Genet. 8 (4), 611-619 (1999)
JOURNAL 99172070
MEDLINE
REMARK Erratum: [[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
REFERENCE 2 (bases 1 to 854)
AUTHORS Brown,A.L. and Kay,G.F.
DIRECT SUBMISSION
JOURNAL Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
FEATURES
Location/Qualifiers
1..854
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/db_xref="dbEST:AA272375"
/chromosomes="X"
/map="near Plp"
/tissue_type="pooled organs"
1..854
/gene="Bex3"
172..546
/gene="Bex3"
/codon_start=1
/product="brain expressed X-linked protein 3"
/protein_id="AAD24431.1"
/db_xref="GI:4580594"
/translation="MANYHOENEEQPLQSQEDRPVGGEGHQAANNNNNNHNN
HNHRRGOARLAPNFRWAI PNMDGLGGDDMMFMEMREIRKLRELQRLNC
LRLLMGLSNHHDHDECLMP"
237 a 212 c 228 g 177 t
```

```
gene
CDS
alignment_scores:
Quality: 567.50 Length: 124
Ratio: 5.113 Gaps: 1
Percent Similarity: 89.516 Percent Identity: 83.065
```

alignment_block:

US-09-327-750D-13 x AF097440

Align seg 1/1 to: AF097440 from: 1 to: 854

1 MetalaAsnIleHisGlnAsnGluMetGluGlnProMetGlnAs 17
|||||

```
172 ATGCCCAATGTCCACCAGGAAACGAAGAGATGGAGCAGCCCTGCAGAA 221
|||||
17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyGlyHisGlnProAlaG 34
|||||
222 TGGACAGGAAGACCGCCTGTGGAGAGAGGTGAGGGCCACACGCTGCTG 271
|||||
34 lyAsn.....ArgArg 37
:||||
272 CAAACAACAACAACAACCAACCAACCAACCAACCAACCAACCAACCA 321
|||||
38 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 54
|||||
322 GGCCAGGCTCGCCACTTGGCCCTTAACCTCCGATGGGCCATTCACCAAC 371
|||||
54 gGlnIleAsnAspGlyMetGlyGlyAspGlyAspMetGluIlePheM 71
|||||
372 GCAGATCAATGACGGGTGGTGGAGATGGAGATGATATGGAAATGTTC 421
|||||
71 etGluGluMetArgGluIleArgArgLysLeuArgGluLeuGlnLeuArg 87
|||||
422 TGGAGGAGATGAGAGATCCGGAGAAAGCTTAGGGAGCTACAGCTGAGA 471
|||||
88 AsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisHisAspHisHi 104
|||||
472 AATTGCTACGATCCTTTATGGGGAGCTGTCTAACCCACCACCATCA 521
|||||
104 sAspGluPheCysLeuMetPro 111
|||||
522 TGATGAATTCGCTTATGCCT 543
|||||
```

seq_name: gb_ro:AF187065

```
seq_documentation_block:
LOCUS AF187065 519 bp mRNA 11-JUN-2000
DEFINITION Rattus norvegicus p75NTR-associated cell death executor (Nade)
mRNA, complete cds.
ACCESSION AF187065
VERSION AF187065.1 GI:8452895
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 519)
AUTHORS Mukai,J., Hachiya,T., Shoji-Hoshino,S., Kimura,M.T., Nadano,D.,
Suvanto,P., Hanaoka,T., Li,Y., Irie,S., Greene,L.A. and Sato,T.A.
NADE, a p75NTR-associated cell death executor, is involved in
signal transduction mediated by the common neurotrophin receptor
p75NTR
J. Biol. Chem. 275 (23), 17566-17570 (2000)
JOURNAL 20298829
MEDLINE 2 (bases 1 to 519)
AUTHORS Mukai,J., Hachiya,T., Hoshino,S., Kimura,M., Nadano,D.,
Hanaoka,T., Li,Y., Irie,S. and Sato,T.
DIRECT SUBMISSION
TITLE Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia
JOURNAL University, 630 West 168th St., P&S II-451, New York, NY 10032, USA
FEATURES
Location/Qualifiers
1..519
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
1..519
/gene="Nade"
44..406
/gene="Nade"
/gene="Nade"
/note="involved in the common neurotrophin receptor
p75NTR-mediated signal transduction; NADE"
/codon_start=1
/product="p75NTR-associated cell death executor"
/protein_id="AAF75130.1"
/db_xref="GI:8452896"
```

```

MEDLINE 20298829
REFERENCE 2 (bases 1 to 891)
AUTHORS Mukai,J., Hachiya,T., Hoshino,S., Kimura,M., Nadano,D., Suvanto,P.,
Hanaoka,T., Li,Y., Irie,S. and Sato,T.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia
University, 630 West 168th St., P&S 11-451, New York, NY 10032, USA
FEATURES
    source
        1..891
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
    gene
        1..891
    CDS
        312..647
        /gene="NADE"
        /note="involved in the common neurotrophin receptor
        /p75NTR-mediated signal transduction"
        /codon_start=1
        /product="p75NTR-associated cell death executor"
        /protein_id="AAF75129.1"
        /db_xref="GI:8452894"
        /translation="NANTHQENEMEQPNQCEEDRPLGGEGHQPAGNRGQARRLA
        PNFRWAIPIRQINDMGDDMEIFMEEMREIRKRLRELQLRNLRLMGLSNHHD
        HHDEFCLMP"
BASE COUNT 251 a 182 c 224 g 234 t
ORIGIN
alignment_scores:
    Quality: 609.00 Length: 111
    Ratio: 5.486 Gaps: 0
    Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
    US-09-327-750D-13 x AF187064
Align seg 1/1 to: AF187064 from: 1 to: 891
1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
312 ATGCCAANTATTCCACGAGAAACGAAGAGATGGAGCGCTATGCAGAA 361
17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyHisGlnProAlaG 34
362 TGGAGAGCAAGACCGCCCTTTGGGAGGAGGTGAAGGCCACCGCTGCAG 411
34 lYAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAla 50
412 GAAATCGACGGGACAGCGCTCGCGACTTGCCTTAATTTTCGATGGGCC 461
51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspPw 67
462 ATACCCCAATAGGCAGATCATATGGATGGGTGGAGATGGAGATGATAT 511
67 tGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGlu 84
512 GGAATAATTTCATGAGGAGATGAGAGAAATCAGAGAAAATCTAGGGAGC 561
84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
562 TGCAGTTGAGGAAATGTCTCGTATCTTATGGGGAGGCTCTCTAATCAC 611
101 HisAspHisHisAspGluPheCysLeuMetPro 111
612 CATGACCATCATGATGAATTTTGCCTTATGCCT 644
seq_name: gb_ro:AF187066
seq_documentation_block:
LOCUS AF187066 700 bp mRNA ROD 11-JUN-2000
DEFINITION Mus musculus p75NTR-associated cell death executor (Nade) mRNA,
complete cds.
ACCESSION AF187066
VERSION AF187066.1 GI:8452897
KEYWORDS house mouse,
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 700)
AUTHORS Mukai,J., Hachiya,T., Shoji-Hoshino,S., Kimura,M.T., Nadano,D.,
Suvanto,P., Hanaoka,T., Li,Y., Irie,S., Greene,L.A. and Sato,T.A.
TITLE Nade, a p75NTR-associated cell death executor, is involved in
signal transduction mediated by the common neurotrophin receptor
p75NTR
JOURNAL J. Biol. Chem. 275 (23), 17566-17570 (2000)
MEDLINE 20298829
REFERENCE 2 (bases 1 to 700)
AUTHORS Mukai,J., Hachiya,T., Hoshino,S., Kimura,M., Nadano,D., Suvanto,P.,
Hanaoka,T., Li,Y., Irie,S. and Sato,T.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia
University, 630 West 168th St., P&S 11-451, New York, NY 10032, USA
FEATURES
    source
        1..700
        /organism="Mus musculus"
        /strain="BALB/c"
        /db_xref="taxon:10090"
    gene
        1..700
    CDS
        177..551
        /gene="Nade"
        /note="involved in the common neurotrophin receptor
        p75NTR-mediated signal transduction; NADE"
        /codon_start=1
        /product="p75NTR-associated cell death executor"
        /protein_id="AAF75131.1"
        /db_xref="GI:8452898"
        /translation="MANYHQENEMEQPLNQCEEDRPLGGEGHQPAGNRGQARRLA
        INHRRGQARRLAPNFRWAIPIRQINDMGDDMEIFMEEMREIRKRLRELQLRNLRL
        RLRLMGLSNHHDHDEFCLMP"
BASE COUNT 178 a 187 c 203 g 132 t
ORIGIN
alignment_scores:
    Quality: 567.50 Length: 124
    Ratio: 5.113 Gaps: 1
    Percent Similarity: 89.516 Percent Identity: 83.065
alignment_block:
    US-09-327-750D-13 x AF187066
Align seg 1/1 to: AF187066 from: 1 to: 700
1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
177 ATGCCCAATGTCACCGAGAAACGAAGAGATGGAGCGCTTGCAGAA 226
17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyHisGlnProAlaG 34
227 TGGACAGGAGAACCGCCCTGTGGGAGGAGGTGAGGGCCACCGCCTGCTG 276
34 lYAsn.....ArgArg 37
277 CAACAACAACAACAACCAACCAACCAACCAACCAACCAACCAACCA 326
38 GlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAsnAr 54
327 GCCCAGGCTCGCGACTTGCCTTAACCTCCGATGGCCATTCCTCAACAG 376
54 gGlnIleAsnAspGlyMetGlyGlyAspGlyAspMetGluIlePheM 71
377 GCAGATGAATGACGGGTGGTGGGATGATGATATGGAATGTTC 426
71 etGluGluMetArgGluIleArgLysLeuArgGluLeuGlnLeuArg 87
```

/translation="MEQPLQNGQEDRPVGGEGHQHPAANNHNNHNNHNNHNNHNNH
RGGARRLAPNFRWAIIPNQMNDGLGDDMEMFMEEMREIRKRLRELQRLNCLRIL
MGELSNHHDHDEFLMP"
BASE COUNT 142 a 130 c 136 g 111 t
ORIGIN

alignment_scores:
Quality: 538.50 Length: 125
Ratio: 5.080 Gaps: 1
Percent Similarity: 84.800 Percent Identity: 79.200

alignment_block:
US-09-327-750D-13 x AF187065 ..

Align seg 1/1 to: AF187065 from: 1 to: 519

6 GlnGluAsnGluGluMetGluGluProMetGlnAsnGlyGluGluAspAr 22
29 CAGGAAACGAGAAATGGAGCAGCCCTGCAGATGGACAGGAGACCG 78
22 gProLeuGlyGlyGluGluHisGlnProAlaGlyAsn..... 35
79 CCCCGTGGGAGGAGGTGAGGGCCACCAGCCCTGCTGCAACACACAAACC 128
36Arg 36
129 ACAACCATACCAACACCAACCAACCAACCAACCAATCATCACCG 178
37 ArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIleProAs 53
179 AGAGGACAGCCCGCGACTTGCCTTAACCTTCGATGGCCATTCCTCCAA 228
53 nArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspMetGluIleP 70
229 CAGCGAGATGAATGATGGGTGGGTGGAGATGATATGGAATGT 278
70 heMetGluGluMetArgGluIleArgArgLysLeuArgGluLeuGlnLeu 86
279 TCATGGAGGAGATGACAGAGATCCGAGAAAGCTTAGGAGCTGCAGTTG 328
87 ArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisAspHI 103
329 AGAATTTGTCGGTATTCTTATGGGGAGCTCTCTAATCACCACGACCA 378
103 sHisAspGluPheCysLeuMetPro 111
379 TCACGATGAATTCGCTTATGCCT 403

seq_name: gb_sts:G72708

seq_documentation_block: 421 bp DNA STS 08-AUG-2001
LOCUS G72708
DEFINITION MARC 4953-4954:991939031:1 SCF - porcine spleen Sus scrofa STS
genomic, sequence tagged site.
ACCESSION G72708
VERSION G72708.1 GI:15146738
KEYWORDS STS.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 421)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L. and
Keeler,J.W.
TITLE Single nucleotide polymorphism (SNP) discovery in expressed porcine
genes
JOURNAL Unpublished (2001)
COMMENT
Contact: Freking BA
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4278

Fax: 402 762 4173
Email: freking@email.marc.usda.gov
Primer A: GCAATGGGTGAACCTCTACT
Primer B: CCAGCAGCAATAGACG
STS size: 500
PCR Profile:

Hotstart: 95 degrees for 15 minutes
Denature: 95 degrees for 30 seconds
Anneal: 56 degrees
Extension: 68 degrees for 2 minutes
Cycles: 32 to 45

Protocol:
Template: 50-200 ng genomic DNA
Primer: each 20 pmoles
dNTPs: each 88 uM
Taq Polymerase: 0.25 units (Qiagen HotStar)

Buffer: Commercially supplied Qiagen HotStar buffer

The STS is derived from PCR amplicons generated from genomic DNA, sequenced from each end using the amplification primers. The sequence does not necessarily represent the entire amplicon. Sequence derived from Polyphred was trimmed from each end of each unique contig until five consecutive bases exceeded a quality score threshold of 20, and the next 10 bases averaged a quality score of 20 or greater. Amplicon size was estimated by agarose gel electrophoresis.

FEATURES
source

Location/Qualifiers
1..421
/organism="Sus scrofa"
/strain="white composite, duroc, meishan, minzhu, fengjing, crossbreds"
/db_xref="taxon:9823"
/sex="male and female"
/clone_lib="SCF - porcine spleen"
/dev_stage="adult"
/note="Organ: spleen"
<1..>421

BASE COUNT 130 a 101 c 86 g 102 t 2 others
ORIGIN

alignment_scores:
Quality: 418.00 Length: 77
Ratio: 5.429 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.701

alignment_block:
US-09-327-750D-13 x G72708/rev ..

Align seg 1/1 to reverse of: G72708 from: 1 to: 421

35 AsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAlaIle 51
421 AATAGACGGGACAGCTGCGGACCTGCCCCCTAATTTCCGATGGCCAT 372
51 eProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspMetG 68
371 ACCCAATAGCGATCAATGATGGGATGGGTGGAGATGGAGATGATGG 322
68 luIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGluLeu 84
321 AATGTTTCATGGAGAGATGAGAGAAATCAGGAGAAAACCTTAGGAGCTG 272
85 GlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHisH 101
271 CAGTTGAGGAATTTGCTCGGTATCCTTATGGGGAGGCTCTCTAATCACC 222
101 sAspHisHisAspGluPheCysLeuMetPro 111
221 TGACCATCATGATGAATTTTGCCTTATGCCT 191

```

seq_name: gb_pr:HSV351F8
seq_documentation_block:
LOGUS HSV351F8 45678 bp DNA PRI 23-NOV-1999
DEFINITION Human DNA sequence from cosmid V351F8, between markers DXS366 and
DXS87 on chromosome X contains ESTs.
ACCESSION 270719
VERSION 270719.1 GI:1261915
KEYWORDS x.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Whiteley,M.
TITLE Direct Submission
JOURNAL Submitted (09-APR-1995) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, UK. E-mail enquires: humquery@sanger.ac.uk
COMMENT IMPORTANT: This sequence is the entire insert of clone V351F8. The
true left end of clone V351F8 is at 1 in this sequence. The true
right end of clone V351F8 is at 45678.
V351F8 is from the human chromosome X-specific cosmid library.
FEATURES
source
location/Qualifiers
1..45678
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="X"
/clone="GHC-351F8"
/clone_lib="SCGV"
1..70
repeat_region
partial
/note="Alu repeat: matches 80. .1 of consensus"
642..930
repeat_region
/note="Alu repeat: matches 1. .308 of consensus"
1406..5179
repeat_region
/note="L1 element fragment"
5243..5276
repeat_region
/note="17 copies of 2 mer 82 % conserved"
5279..5567
repeat_region
partial
/note="Alu repeat: matches 304. .1 of consensus"
9975..10369
repeat_region
/note="MSTA element fragment"
10399..12017
repeat_region
/note="MSTAR element fragment"
11318..11978
repeat_region
/note="THE1B element fragment"
11715..11846
repeat_region
/note="MLTIR element fragment"
12074..12170
repeat_region
/note="MSTA element fragment"
12227..12486
repeat_region
/note="MSTA element fragment"
12437..12502
repeat_region
/note="MSTC element fragment"
12756..13059
repeat_region
partial
/note="Alu repeat: matches 308. .1 of consensus"
14478..14650
repeat_region
/note="THE1B element fragment"
14491..14536
repeat_region
/note="MSTA element fragment"
14715..14831
repeat_region
/note="THE1B element fragment"
15099..15533
repeat_region
/note="L1 element fragment"
15534..15626
repeat_region
/note="31 copies of 3 mer 85 % conserved"
15566..15619
repeat_region
/note="3 copies of 18 mer 98 % conserved"
16352..16436
repeat_region
/note="L1 element fragment"
16728..16763
repeat_region
/note="9 copies of 4 mer 94 % conserved"
16764..17054
repeat_region
partial
/note="Alu repeat: matches 308. .1 of consensus"
19420..19449
repeat_region
/note="15 copies of 2 mer 87 % conserved"
complement(19537..20013)
misc_feature
/note="match: 3' EST N51315 clone 283089"
20071..20114
repeat_region
/note="22 copies of 2 mer 98 % conserved"
20073..20112
repeat_region
/note="10 copies of 4 mer 100 % conserved"
20073..20126
repeat_region
/note="3 copies of 18 mer 87 % conserved"
23088..23300
repeat_region
/note="L1 element fragment"
23311..23490
repeat_region
/note="MLTIA element fragment"
23318..23489
repeat_region
/note="MLTIB element fragment"
23362..23491
repeat_region
/note="MLTIC element fragment"
23507..24407
repeat_region
/note="L1 element fragment"
24376..24462
repeat_region
/note="MSTC element fragment"
24379..24491
repeat_region
/note="MSTA element fragment"
24688..24738
repeat_region
/note="MSTA element fragment"
24810..25016
repeat_region
/note="L1 element fragment"
25254..25421
repeat_region
/note="L1 element fragment"
25439..25696
repeat_region
/note="MER25 element fragment"
26284..26343
repeat_region
/note="MLTID element fragment"
26428..26491
repeat_region
/note="MLTIB element fragment"
26465..26518
repeat_region
/note="MLTIE element fragment"
27446..27667
repeat_region
/note="L1 element fragment"
28954..29247
repeat_region
partial
/note="Alu repeat: matches 308. .1 of consensus"
30160..30325
repeat_region
/note="2 copies of 83 mer 98 % conserved"
32588..32875
repeat_region
/note="match: 5' EST H68599 clone 239077"
complement(32825..33230)
misc_feature
/note="match: 3' EST H68239 clone 289077"
35551..35802
repeat_region
partial
/note="Alu repeat: matches 1. .260 of consensus"
35804..35839
repeat_region
/note="18 copies of 2 mer 83 % conserved"
38625..38705
repeat_region
/note="MLTIR element fragment"
38747..38839
repeat_region
/note="MLTIR element fragment"
38949..39032
repeat_region
/note="MLTIE element fragment"
38965..39076
repeat_region
/note="MLTID element fragment"
BASE COUNT 14260 a 9135 c 9327 g 12956 t
ORIGIN
alignment_scores: Quality: 300.00 Length: 114

```


COMMENT

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX>

035619 IS FROM THE LIBRARY RPCI4 CONSTRUCTED AT THE ROSWELL PARK Cancer Institute by the group of Pieter de Jong. For further

details see <http://bacpac.med.utoronto.edu/vector/pclpAC2>

IMPORTANT: This sequence is not the entire insert of clone 635G19. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between

The true left end of clone U65A4 (Z81014) is at 69549 in this sequence. The true right end of clone U101D3 (Z85997) is at 100 in this sequence.

```

location/Qualifiers
1. .69648
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

```
/chrChromosome="X"
/map="q22.1-22.3"
/clone="RP4-635G19"
/clone_lib="RPCI-4"
```

```

region
    /note="TAR1 repeat: matches 707. .749 of consensus"
    501. .661
region
    /note="AluJb repeat: matches 137. .311 of consensus"
    455. .749

```

```

/note="118 copies 2 mer aa 70% conserved"
915. 1331
region
/note="NSTA repeat: matches 1. 419 of consensus"

```

```

region
/note="92 copies 2 mer ga 77% conserved"
1927. 2020
/note="MIR repeat: matches 109. 217 of consensus"

```

```

2004.12.02
region:
/note="L1ME2 repeat: matches 5825. .6022 of consensus"

```

region	2405..2574	/note="Alu repeat: matches 1..313 of consensus"
region	2575..2819	/note="L1ME2 repeat: matches 5583..5825 of consensus"
region	2821..3039	/note="Alu repeat: matches 1..311 of consensus"
region	3104..3384	

region	/note="l1m2 repeat: matches 5304. .5591 of consensus"
region	3391. .4007
region	/note="l1m1 repeat: matches 5558. .6162 of consensus"
region	4012. .4137

	/note="Higger2a repeat: matches 11. .123 of consensus"
I-region	4141..4252
region	/note="I1 repeat: matches 2920. .3031 of consensus"
	4265..4350
region	/note="HY4 repeat: matches 1. .87 of consensus"

region	4500..4670	/note="MER74A repeat: matches 28. .369 of consensus"
	4914..5095	
_region		
	5224..5555	/note="MER5A repeat: matches 8. .189 of consensus"
region		

Location/Qualifiers
1. .69648

```

1. 59648
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q22.1-22.3"
/clone="RP4-635G19"
/clone_lib="RPCI-4"

```

```

439. .483
/NOTE="PAR1 repeat: matches 707. .749 of consensus"
501. .661
/NOTE="AluDb repeat: matches 137. .311 of consensus"
562. .897
/NOTE="118 copies 2 mer aa 70% conserved"
915. .1331
/NOTE="NSTA repeat: matches 1. .419 of consensus"
1334. .1517
/NOTE="92 copies 2 mer ga 77% conserved"
1927. .2020
/NOTE="MIR repeat: matches 109. .217 of consensus"
2066. .2262
/NOTE="L1ME2 repeat: matches 5825. .6022 of consensus"
2263. .2574
/NOTE="AluSp repeat: matches 1. .313 of consensus"
2375. .2819
/NOTE="L1ME2 repeat: matches 5583. .5825 of consensus"
2821. .3039
/NOTE="Alu repeat: matches 1. .311 of consensus"
3104. .3384
/NOTE="L1ME2 repeat: matches 5304. .5591 of consensus"
3391. .4007
/NOTE="L1M61 repeat: matches 5558. .6162 of consensus"
4012. .4137
/NOTE="Trigger2a repeat: matches 11. .123 of consensus"
4141. .4252
/NOTE="L1 repeat: matches 2920. .3031 of consensus"
4265. .4350
/NOTE="HY4 repeat: matches 1. .87 of consensus"
4350. .4678
/NOTE="MER74A repeat: matches 28. .369 of consensus"
4914. .5095
/NOTE="MER5A repeat: matches 8. .189 of consensus"
5224. .5555

```


seq_name: gb_sts:G35294

seq_documentation_block: 477 bp DNA STS 02-OCT-1997
 LOCUS G35294 human STS SHGC-37409, sequence tagged site.
 DEFINITION G35294
 ACCESSION G35294.1 GI:2459462
 VERSION STS: STS sequence; primer; sequence tagged site.
 KEYWORDS STS: STS sequence; primer; sequence tagged site.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 477)
 AUTHORS Myers,R.M.
 JOURNAL Unpublished (1997)

Contact: Richard M. Myers
 Stanford Human Genome Center (SHGC)
 Stanford University School of Medicine
 Department of Genetics, M-344, Stanford, CA 94305, USA
 Tel: 4157259687
 Fax: 4157259689
 Email: myers@shgc.stanford.edu

Primer A: AACATCTTTCATGAAAGTTGATG
 Primer B: CTTTGGCATCTTCTCGAA
 STS size: 106
 PCR Profile:

Initial incubation: 95 degrees C for 10 minutes

Denaturation: 94 degrees C for 15 seconds
 Annealing: 60 degrees C for 30 seconds
 Polymerization: 72 degrees C for 23 seconds
 PCR Cycles: 30
 Thermal Cycler: Perkin Elmer 9600

Protocol:
 Template: 25 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 AmpliTaq Gold Polymerase: 0.07 units/ul
 Total Vol: 5 ul

Buffer:
 MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from N51315
 -- Washington University/Merck EST sequence.

FEATURES
 Location/Qualifiers

1..477

/organism="Homo sapiens"

/db_xref="taxon:9606"

44..149

44..67

primer_bind complement(130..149)

primer_bind 153 a 115 c 79 g 130 t

BASE COUNT

ORIGIN

alignment_scores:
 Quality: 249.50 Length: 83
 Ratio: 3.960 Gaps: 2
 Percent Similarity: 75.904 Percent Identity: 61.446

alignment_block:

US-09-327-750D-13 x G35294/rev ..

Align seg 1/1 to reverse of: G35294 from: 1 to: 477

31 GlnProAlaGlyAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPh 47

```

:::|||||:::|||||:::|||||
471 GAGCCTGGAGGAATGTTAAAGGGTTTGGGCTCCACCTGCCCGGGTTT 422
|
47 eArgTrpAlaIleProAsnArgGlnIleAsn.....AspGlyMetGlyG 62
|   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
421 TGGAGAGGATGTGCCCAATAGGCTTGTGCGATAACATTTGATATAGATG 372
|   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
62 lYAspGlyAspAspMetGluIlePheMetGluGluMetArgGluIleArg 78
|   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
371 GAGATGGAGATGATATGGAACGGTTTCATGGAGGAGATGAGAGCTAAGG 322
|   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
79 ArgLysLeuArgGluLeuGlnLeuArgAsnCysLeuArgIleLeuMetG1 95
|   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
321 AGCAAAATTAGGAACTTCAGTTGAGGTACAGCTCGGCATCTCTATAGG 272
|   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
95 yGluLeuSerAsnHisAspHisAspGluPheCysLeuMetPro 111
|   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
271 GGAC...CTCCTCACCATCATCATGATGAGTTTTCGCTTATGCTT 226
|   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
seq_name: gb_pr:HS714B7

```

seq_documentation_block:

LOCUS HS714B7 98274 bp DNA PRI 12-DEC-1999
 DEFINITION Human DNA sequence from clone CTA-714B7 on chromosome 22q12.2-13.2
 Contains pseudogene similar to part of COX7B (cytochrome c oxidase subunit VIIb), a novel pseudogene, ESTs, STS and GSSs, complete sequence.
 ACCESSION Z99755
 VERSION Z99755.1 GI:3036782
 KEYWORDS HTG; COX7B; cytochrome c oxidase.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 98274)
 DIRECT SUBMISSION Connor,R.
 REFERENCE Direct Submission
 AUTHORS Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,
 JOURNAL CB10 ISA, UK. E-mail enquiries: humquery@sanger.ac.uk
 COMMENT requests: clonerequest@sanger.ac.uk
 On Apr 8, 1998 this sequence version replaced gi:2578134.
 This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome 22 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr22>
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep CTA-714B7 is from the human BAC library described in U-J. Kim et al. (1996)

Genomics 34, 213-218.

VECTOR: pBelOBAC11

This sequence is the entire insert of clone CTA-714B7.

FEATURES

source

Location/Qualifiers

1..98274

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="22"

/map="q12.2-13.2"

/clone="CTA-714B7"


```
repeat_region /clone_lib="CIT978SK-A2"
repeat_region /note="14 copies 2 mer ac 100 conserved"
repeat_region /note="L2 repeat: matches 7840. .7913 of consensus"
misc_feature 1177. .1446
repeat_region /note="match: GSS: Em:AQ373724"
repeat_region /note="L2 repeat: matches 2693. .2750 of consensus"
repeat_region 1301. .1356
repeat_region 2568. .5038
repeat_region /note="L1MD1 repeat: matches 3673. .6170 of consensus"
repeat_region 5036. .5486
repeat_region /note="L1MCE repeat: matches 2052. .2207 of consensus"
repeat_region 5496. .6885
repeat_region /note="L1MD repeat: matches -1. .1538 of consensus"
repeat_region 6896. .7185
misc_feature complement(7187. .7611)
repeat_region /note="match: GSS: Em:AQ566001"
repeat_region 7369. .7594
repeat_region /note="MIR repeat: matches 34. .256 of consensus"
misc_feature complement(7713. .8092)
repeat_region /note="match: GSS: Em:AQ071449"
repeat_region 7904. .7978
repeat_region /note="MIR repeat: matches 20. .95 of consensus"
repeat_region 8086. .8300
misc_feature 8154. .8565
repeat_region /note="match: GSS: Em:AQ059599"
repeat_region 8349. .8864
misc_feature /note="AluSx repeat: matches 1. .280 of consensus"
repeat_region complement(8897. .9290)
repeat_region /note="match: GSS: Em:AQ343214"
misc_feature 9235. .9361
repeat_region /note="L2 repeat: matches 2619. .2749 of consensus"
repeat_region 9378. .9890
repeat_region /note="match: GSS: Em:AQ284832"
repeat_region 10135. .10428
repeat_region /note="AluSx repeat: matches 1. .294 of consensus"
repeat_region 10429. .10563
repeat_region /note="AluSg/x repeat: matches 1. .135 of consensus"
repeat_region 10811. .10868
repeat_region /note="L2 repeat: matches 2648. .2705 of consensus"
repeat_region 10819. .11064
repeat_region /note="MIR repeat: matches 2. .255 of consensus"
repeat_region 11112. .11324
repeat_region /note="MIR repeat: matches 11. .255 of consensus"
repeat_region 11869. .12012
repeat_region /note="MIR repeat: matches 20. .167 of consensus"
repeat_region 12925. .13223
repeat_region /note="AluY repeat: matches 1. .300 of consensus"
repeat_region 13290. .13495
misc_feature /note="MER58A repeat: matches 9. .214 of consensus"
repeat_region 13338. .13833
repeat_region /note="match: GSS: Em:AQ568864"
repeat_region 13688. .13749
repeat_region 13840. .13966
repeat_region /note="MIR repeat: matches 77. .139 of consensus"
repeat_region 14162. .14245
repeat_region /note="MIR repeat: matches 22. .147 of consensus"
repeat_region 14925. .15058
repeat_region /note="MIR repeat: matches 46. .136 of consensus"
repeat_region 15994. .16040
repeat_region /note="MIR repeat: matches 48. .191 of consensus"
repeat_region 16116. .16184
repeat_region /note="MIR repeat: matches 102. .144 of consensus"
repeat_region 16897. .17005
repeat_region /note="MER5B repeat: matches 3. .70 of consensus"
repeat_region 17618. .17700
repeat_region /note="MIR repeat: matches 9. .118 of consensus"
repeat_region /note="MIR repeat: matches 108. .192 of consensus"
repeat_region 17748. .17912
repeat_region /note="MIR repeat: matches 73. .245 of consensus"
```

```
repeat_region 18657. .18871
repeat_region /note="MIR repeat: matches 35. .252 of consensus"
repeat_region 18927. .19230
repeat_region /note="AluSx repeat: matches 1. .301 of consensus"
repeat_region 19492. .19559
repeat_region /note="MIR repeat: matches 2. .67 of consensus"
repeat_region 19560. .19821
repeat_region /note="AluY repeat: matches 34. .295 of consensus"
repeat_region 19822. .19959
repeat_region /note="MIR repeat: matches 67. .212 of consensus"
repeat_region 19987. .20286
repeat_region /note="AluY repeat: matches 1. .300 of consensus"
repeat_region 20316. .20512
repeat_region /note="MIR repeat: matches 82. .262 of consensus"
repeat_region 20513. .20741
repeat_region /note="L1MD1 repeat: matches 5970. .6224 of consensus"
repeat_region 20769. .21074
repeat_region /note="AluSg1 repeat: matches 1. .303 of consensus"
repeat_region 21085. .21207
repeat_region /note="L1M2 repeat: matches 1776. .1898 of consensus"
repeat_region 21212. .21391
repeat_region /note="L1MD repeat: matches -6. .178 of consensus"
repeat_region 21392. .21471
repeat_region /note="MIR repeat: matches 3. .86 of consensus"
repeat_region 22146. .22441
repeat_region /note="AluSx repeat: matches 1. .296 of consensus"
repeat_region 22526. .22610
repeat_region /note="L1M2 repeat: matches 267. .349 of consensus"
repeat_region 22940. .23204
repeat_region /note="HERV16 repeat: matches 378. .650 of consensus"
repeat_region 23205. .23514
repeat_region /note="AluSg repeat: matches 1. .311 of consensus"
repeat_region 23919. .24194
repeat_region /note="AluSx repeat: matches 36. .311 of consensus"
repeat_region 24496. .24590
repeat_region /note="L1M2 repeat: matches 253. .349 of consensus"
repeat_region 24871. .24916
repeat_region /note="MIR repeat: matches 102. .143 of consensus"
repeat_region 25378. .25555
repeat_region /note="L2 repeat: matches 2330. .2519 of consensus"
repeat_region 25673. .25910
repeat_region /note="MIR repeat: matches 4. .255 of consensus"
repeat_region 26722. .26798
repeat_region /note="L2 repeat: matches 2572. .2649 of consensus"
repeat_region 27040. .27107
repeat_region /note="L2 repeat: matches 2489. .2500 of consensus"
repeat_region 27108. .27374
repeat_region /note="AluY repeat: matches 41. .304 of consensus"
repeat_region 27375. .27471
repeat_region /note="L2 repeat: matches 2372. .2489 of consensus"
repeat_region 27651. .27945
repeat_region /note="AluSx repeat: matches 1. .299 of consensus"
repeat_region 28235. .28876
repeat_region /note="L2 repeat: matches 1113. .1817 of consensus"
repeat_region 28975. .29097
repeat_region /note="MER96 repeat: matches 42. .173 of consensus"
repeat_region 29173. .29196
misc_feature /note="L2 copies 2 mer aa 100 conserved"
repeat_region 29832. .30260
repeat_region /note="match: GSS: Em:AQ075670"
repeat_region 30392. .30597
repeat_region /note="MIR repeat: matches 35. .252 of consensus"
repeat_region 30845. .31010
repeat_region /note="MER91B repeat: matches 2. .162 of consensus"
repeat_region 31056. .31094
repeat_region /note="MIR repeat: matches 218. .256 of consensus"
repeat_region 31728. .31791
repeat_region /note="L2 copies 4 mer tgg 76 conserved"
repeat_region 31858. .31952
misc_feature /note="MIR repeat: matches 48. .153 of consensus"
repeat_region complement(33437. .33589)
repeat_region /note="match: STS: Em:H55731"
repeat_region 33543. .33633
```

```
/note="MIR repeat: matches 84...165 of consensus"
33785..33966
/note="MER91A repeat: matches 1..185 of consensus"

alignment_scores:
  Quality: 241.50      Length: 67
  Ratio: 4.093         Gaps: 2
  Percent Similarity: 88.060  Percent Identity: 77.612

alignment_block:
US-09-327-750D-13 x HS714B7 ..
Align seg 1/1 to: HS714B7 from: 1 to: 98274

43 LeuAlaProAsnPheArgTirAlaIleProAsnArgGlnIleAsnAspGI 59
|||||
62887 TTGGCCCTTAATTTTCAATGGCCATACCAATAGGAGGTCAGATGG 62936
|||||
59 yMetGlyGlyAspGlyAspMetGluIlePheMetGluGluMetArgG 76
|||||
62937 GATGAATAGAGTGAGATGATATGAAATGTTTCATGAGGAGATGAGAG 62986
|||||
76 luIleArgArg-LysLeuArgGluLeuGlnLeuArgAsnCysLeuArgII 92
|||||
62987 GAATCAGGAGAGAAT...AAGGAGCTACAATTGAGGAATTTCTGTGTAT 63033
|||||
92 eLeuMetGlyGluLeuSerAsnHisHisAspHisHisAspGluPheCys 108
|||||
63034 CCTATATGGAGAG.CHTGTATATCCCATGACCATCATGATGAATTTTGC 63081

seq_name: gb_pr:AF183416

seq_documentation_block:
LOCUS AF183416 642 bp mRNA PRI 02-SEP-2000
DEFINITION Homo sapiens ovarian granulosa cell 13.0 kDa protein HGR74 homolog
mRNA, complete cds.
ACCESSION AF183416
VERSION AF183416
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Peng,Y., Qian,b., Tu,Y., Xu,S., Han,Z., Fu,G. and Chen,Z.
TITLE A novel gene expressed in human adrenal gland
JOURNAL Unpublished
REFERENCE
AUTHORS Peng,Y., GU,W., Huang,C., Xu,S., Han,Z., Fu,G. and Chen,Z.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-1999) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China
FEATURES
source
1..642
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="adrenal gland"
3..380
/note="HGR74-h protein"
/codon_start=1
/evidence="not_experimental"
/product="ovarian granulosa cell 13.0 kDa protein HGR74
homolog"
/db_xref="GI:9963771"
/translation="MESKEKRAVNSLSMENAOEKEOVANKGEPLALPLDAGEYC
VPRGNRRFRVROPILQYRWDMMHRLGEPQARMREENMERIGEEVRLMEKLRKQLS
HSLRAVSTDPDHPHDDHDECLMP"

BASE COUNT 204 a 118 c 157 g 163 t
ORIGIN
```

```
alignment_scores:
  Quality: 194.50      Length: 113
  Ratio: 2.739         Gaps: 3
  Percent Similarity: 62.832  Percent Identity: 43.363

alignment_block:
US-09-327-750D-13 x AF183416 ..
Align seg 1/1 to: AF183416 from: 1 to: 642

1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
|||
42 ATGAAATAATGCCAACCAAGAAATAATGAAGAAAGGAGCAAGTTGCTAATAA 91
|||||
17 nGlyGluGlu...AsparGProLeuGlyGlyGlyGluGlyHisGlnProA 33
|||||
92 AGGGAGAGCCCTTGGCCCTTGGATGCTGGTGAATACTGTGTGCTA 141
|||||
33 laGlyAsnArgArgGly...GlnAlaArgArgLeuAlaProAsnPheArg 48
|||||
142 GAGAAATCTAGGCGGTTCGCCGTTAGGAGCCCATCTCCAGTATAGA 191
|||||
49 TrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAs 65
|||
192 TGGCATATGATGATAGCTTGGAGAACCCACAGGAGGATGAGAGAAGA 241
|||||
65 pasMetGluIlePheMetGluGluMetArgGluIleArgArgLysLeuA 82
|||||
242 GAATATGAAAGGATTGGGGAGGAGGTGACACAGCTGATGGAAAGCTGA 291
|||||
82 rgGluLeuGlnLeuArgAsnCysLeuIleLeuMetGlyGluLeuSer 98
|||||
292 GGGAAACAGCTTACGTCATAGTCTGGGGCAGTCAGCAGCTGAC...CCC 338
|||||
99 AsnHisHisAspHisHisAspGluPheCysLeuMetPro 111
|||||
339 CCTCACCATGACCATCATGATGAGTTTTCCTTTATGCCCC 377

seq_name: gb_pr:AF220189

seq_documentation_block:
LOCUS AF220189 828 bp mRNA PRI 04-MAY-2000
DEFINITION Homo sapiens uncharacterized hypothalamus protein HBEX2 mRNA,
complete cds.
ACCESSION AF220189
VERSION AF220189
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Xiao,H., Song,H., Gao,G., Ren,S., Chen,Z. and Han,Z.
TITLE A novel gene expressed in human hypothalamus
JOURNAL Unpublished
REFERENCE
AUTHORS Xiao,H., Song,H., Gao,G., Ren,S., Chen,Z. and Han,Z.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China
FEATURES
source
1..828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="hypothalamus"
167..544
/codon_start=1
/product="uncharacterized hypothalamus protein HBEX2"
/db_xref="GI:7689029"
/translation="MESKEKRAVNSLSMENAOEKEOVANKGEPLALPLDAGEYC
VPRGNRRFRVROPILQYRWDMMHRLGEPQARMREENMERIGEEVRLMEKLRKQLS
```

```
HSLRAVSTDPHPHHDHDFCLMP"
BASE COUNT      256 a      170 c      221 g      181 t
ORIGIN

alignment_scores:
  Quality: 194.50      Length: 113
  Ratio: 2.739        Gaps: 3
  Percent Similarity: 62.832      Percent Identity: 43.363

alignment_block:
US-09-327-750D-13 x AF220189  ..

Align seg 1/1 to: AF220189 from: 1 to: 828

1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
   ||| ||| :|||:|||||:||||| :|||: :|||:
206 ATGGAATGCCAACCAAGAAATGAAGAAAGGAGCAAGTGTCTAATAA 255
   ||| |||:|||||:||||| :|||: :|||:
17 nGlyGluGlu...AspArgProLeuGlyGlyGlyHisGlnProA 33
   :||| ||| ||| ||| :||| |||
256 AGGGAGCCCTTGGCCCTCTTGGATGCTGGTGAATACTGTGCTA 305
   :||| ||| ||| ||| :||| |||
33 laGlyAsnArgArgGly...GlnAlaArgLeuAlaProAsnPheArg 48
   ||| ||| ||| ||| :||| |||
306 GAGGAATCGTAGCGGTTCGCGTTAGCGAGCCCATCTGCGAGTATAGA 355
   ||| ||| ||| ||| :||| |||
49 TrpAlaIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAs 65
   ||| ||| :||| ||| :||| |||
356 TGGATATGATGATAGGCTTGGAGACACAGGCAAGGATGAGAGAAGA 405
   :||| ||| ||| ||| :||| |||
65 pAspMetGluIlePheMetGluMetArgGluIleArgArgLysLeuA 82
   :||| ||| ||| ||| :||| |||
406 GAATATGGAAGGATTGGGAGGAGGTGACACAGCTGATGGAAGAGCTGA 455
   ||| ||| ||| ||| :||| |||
82 rgGluLeuGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSer 98
   ||| ||| ||| ||| :||| |||
456 GGGAAAGCAGTTGAGTCATAGTCTGGCGGAGTCAGCACTGAC...CCC 502
   ||| ||| ||| ||| :||| |||
99 AsnHisHisAspHisHisAspGluPheCysLeuMetPro 111
   ||| ||| ||| ||| :||| |||
503 CCTCACCATGACCATGATGAGTTTGGCTTATGCC 541
```

THIS PAGE BLANK (USPTO)

OM of: US-09-327-750D-13 to: EST:* out_format : pfs
Date: Mar 11, 2002 2:16 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE1=frame+ p2n model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US09327750/runat_11032002_101153_20308/app_query.fasta_1.1472
-DB-EST -QFMT=fastap -SUFFIX=rest -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCI=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FCAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blotum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500
-MINLEN=0 -MAXLEN=2000000000 -USER=US09327750.@CGNL_1_5654
-NCPU=6 -TCPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-13
Query length: 111
Database: EST:*
Database sequences: 11351937
Database length: 1077921985
Search time (sec): 4085.940000

score_list:

Sequence	Strtd	Orig	ZScore	EScore	Len	Documentation
gb_est2:N34237	+	609.00	1048.29	3.2e-49	519	N34237.yx79c03.r1 Soares melanc
gb_est2:BG654378	+	609.00	1048.04	3.3e-49	535	BG654378.lb39602.y1 HR85 islet
gb_est1:AW512400	-	609.00	1046.90	3.8e-49	612	AW512400.xx74b09.x1 NCI_CGAP_Ly
gb_est1:AI1193112	-	609.00	1046.86	3.8e-49	615	AI1193112.qe69q09.x1 Soares.feta
gb_est1:AA576958	-	609.00	1046.79	3.8e-49	620	AA576958.nm82804.s1 NCI_CGAP_C
gb_est1:AI929106	-	609.00	1046.24	4.1e-49	622	AI929106.au65b10.y1 Schneider
gb_est2:BG745809	-	609.00	1046.08	4.2e-49	674	BG745809.602723951t1 NIH_MGC_11
gb_est1:AV701506	+	609.00	1045.53	4.5e-49	720	AV701506.AV701506.ADB Homo sapi
gb_est2:BF058872	+	600.00	1031.77	2.6e-47	586	BF058872.7k35g02.x1 NCI_CGAP_Ov
gb_est2:BG482054	+	597.00	1027.03	4.8e-48	557	BG482054.602527934f1 NIH_MGC_21
gb_est2:BG705685	+	597.00	1026.88	4.9e-48	567	BG705685.602668910f1 NIH_MGC_96
gb_est2:BE96163	+	597.00	1025.39	6.0e-48	676	BE96163.601438756f1 NIH_MGC_72
gb_est2:BG933540	+	596.00	1025.42	5.9e-48	550	BG933540.PM0-HT0911-080201-012
gb_est1:BE52272	+	591.00	1016.67	1.8e-47	559	BE522722.60114115f1 NIH_MGC_16
gb_est1:AI928835	-	589.00	1012.27	3.2e-47	626	AI928835.au65b10.y1 Schneider
gb_est1:AA514324	-	585.00	1005.69	7.5e-47	603	AA514324.nf56b04.s1 NCI_CGAP_C
gb_est2:BG994124	+	584.00	1005.92	7.2e-47	479	BG994124.PM0-HT0911-120201-015
gb_est2:BG764807	+	584.00	1001.81	1.2e-46	779	BG764807.602737223f1 NIH_MGC_49
gb_est2:W25163	+	583.00	1004.69	8.5e-47	452	W25163.zb69h06.r1 Soares.fetal
gb_est2:W80994	+	583.00	1000.81	1.4e-46	715	W80994.zd84f08.r1 Soares.fetal
gb_est1:BE709912	-	582.00	1003.41	1.0e-46	429	BE709912.IL3-HT0619-280600-192
gb_est1:BE378166	-	582.00	1002.15	1.2e-46	498	BE378166.601237907f1 NIH_MGC_44
gb_est2:W40428	+	582.00	1001.99	1.2e-46	507	W40428.zb74a05.r1 Soares.fetal
gb_est2:BE935913	-	582.00	1001.68	1.2e-46	526	BE935913.OV2-NN0053-160800-314
gb_est2:BF198483	+	582.00	1001.67	1.3e-46	527	BF198483.248426.MARC.2P1G.Sus
gb_est1:BE265852	+	582.00	1001.43	1.3e-46	542	BE265852.60193356f1 NIH_MGC_7
gb_est1:AW131584	-	582.00	1001.03	1.4e-46	568	AW131584.xf31g01.x1 NCI_CGAP_Bi
gb_est2:BG737381	+	582.00	1000.96	1.4e-46	573	BG737381.602661455f1 NIH_MGC_21
gb_est1:AI075989	-	582.00	1000.94	1.4e-46	574	AI075989.ov47b05.x1 Soares.test
gb_est2:BF316261	+	582.00	1000.74	1.4e-46	588	BF316261.601899787f1 NIH_MGC_19
gb_est2:BG715325	+	582.00	1000.57	1.4e-46	600	BG715325.602677333f1 NIH_MGC_96
gb_est1:AV709328	+	582.00	1000.43	1.5e-46	610	AV709328.AV709328.ADC Homo sapi
gb_est2:BG705415	+	582.00	1000.36	1.5e-46	615	BG705415.602685538f1 NIH_MGC_95
gb_est2:BF725090	+	582.00	999.81	1.6e-46	656	BF725090.bx12a12.y1 Human iris
gb_est1:AV752634	+	582.00	999.70	1.6e-46	665	AV752634.AV752634.NPD Homo sapi
gb_est2:BF183092	+	582.00	999.65	1.6e-46	669	BF183092.601809666f1 NIH_MGC_18
gb_est2:BF032616	+	582.00	999.54	1.6e-46	678	BF032616.601453033f1 NIH_MGC_66
gb_est1:BE727430	+	582.00	999.49	1.7e-46	682	BE727430.601560632f1 NIH_MGC_20
gb_est2:BE989301	+	582.00	999.45	1.7e-46	685	BE989301.601681102f1 NIH_MGC_9
gb_est1:BE729899	+	582.00	999.44	1.7e-46	686	BE729899.601564946f1 NIH_MGC_20

gb_est2:BG825224 + 582.00 999.39 1.7e-46 690 | BG825224 602747641f1 NIH_MGC
gb_est2:BI114083 + 582.00 999.35 1.7e-46 693 | BI114083 602861185f2 NIH_MGC
gb_est1:AW161416 + 582.00 999.30 1.7e-46 697 | AW161416 au81e08.y1 Schneider
gb_est1:BE385584 + 582.00 999.28 1.7e-46 699 | BE385584 601275870f1 NIH_MGC
gb_est2:BF970435 + 582.00 999.25 1.7e-46 701 | BF970435 602272771f1 NIH_MGC
seq_name: gb_est2:N34237

seq_documentation_block:
LOCUS N34237 519 bp mRNA EST 16-JAN-1996
DEFINITION YX79C03.r1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
IMAGE:267940 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0 KD
PROTEIN HGR74 (HUMAN); mRNA sequence.
ACCESSION N34237
VERSION N34237.1 GI:1155379
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
R., Williamson,A., Wohlmann,P. and Wilson,R.
The WashU-Merck EST Project
Contacted: (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 449
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: T7
High quality sequence stop: 449.
Location/Qualifiers
1..519
/organism="Homo sapiens"
/db_xref="GDB:3877582"
/db_xref="taxon:9606"
/clone="IMAGE:267940"
/clone_lib="Soares melanocyte 2NbHM"
/sex="Male"
/tissue_type="melanocyte"
/lab_host="DH10B (ampicillin resistant)"
/note="vector: pT73B (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTCGAGCGCCGAGTGTGTGTGTGTGTGTGTGTGT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo. RNA from normal foreskin melanocytes
(FS374) was kindly provided by Dr. Anthony P. Albino."

alignment_scores:
Quality: 609.00 Length: 111
Ratio: 5.486 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-327-750D-13 x N34237
Align seg 1/1 to: N34237 from: 1 to: 519

BASE COUNT 148 a 114 c 153 g 98 t
ORIGIN

alignment_scores:
Quality: 609.00 Length: 111
Ratio: 5.486 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-327-750D-13 x N34237
Align seg 1/1 to: N34237 from: 1 to: 519

```

1 MetAlaAsnIleHisGlnGluAsnGluGluMetGluGlnProMetGlnAs 17
|||||
166 ATGGCAAAATATTCACCAAGAAACGAAGAGATGGAGCAGCCTATGCAGAA 215
|||||
17 nGlyGluGluAsnArgProLeuGlyGlyGlyGlyGlyGlyHisGlnProAlaG 34
|||||
216 TGGAGAGGAAGACCGCCCTTTGGAGAGAGTGGAAGCCACACGCTGCAG 265
|||||
34 LyAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTTPAla 50
|||||
266 GAAATCGACGGGACAGAGCTGCGGACTTGCCCTTAATTTTCGATGGGCC 315
|||||
51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspAspMe 67
|||||
316 ATACCAATAGGCAGATCAATGATGGATGGATGGGTGGAGATGGAGATGATAT 365
|||||
67 tGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGluL 84
|||||
366 GGAATATATTCATGGAGAGATGAGAGAAATCAGAAGAAACTTAGGGAGC 415
|||||
84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
|||||
416 TGCAGTTGAGGAATGTCTGCGTATCTTATGGGGAGCTCTCTAATCAC 465
|||||
101 HisAspHisHisAspGluPheCysLeuMetPro 111
|||||
466 CATGACCATCATCATCAATTTTGCCTTATGCCT 498
|||||

```

seq_name: gb_est2:BG654378

```

seq_documentation_block:
LOCUS      BG654378          535 bp      mRNA          05-JUL-2001
DEFINITION ib39e02.y1 HR85 islet Homo sapiens cDNA 5' similar to SM:HG74_HUMAN
            O00994 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74. ;, mRNA
            sequence.

```

```

ACCESSION  BG654378
VERSION     BG654378.1  GI:13791787
KEYWORDS    EST.
SOURCE      human.

```

ORGANISM

```

Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE

```

AUTHORS
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
,M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R., Williams,T.,
Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: ib39e02.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@bioh.p.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 422.
Location/Qualifiers
1..535
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"

```

TITLE

JOURNAL

COMMENT

seq_name: gb_est1:AW512400

seq_documentation_block:

```

LOCUS      AW512400          612 bp      mRNA          03-MAR-2000
DEFINITION xx74b09.x1 NC1_CGAP_Lym12 Homo sapiens cDNA clone IMAGE:2849369 3'
            similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74
            (HUMAN);, mRNA sequence.

```

ACCESSION AW512400

VERSION AW512400.1 GI:7150478

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE

1 (bases 1 to 612)

NCT-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Life Technologies catalog #: 11547-015

/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1: NotI; Site_2: XhoI; cDNA made by oligo-dT priming. -lkb. 5' Size selected on agarose gel. Average insert size -lkb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

BASE COUNT 153 a 121 c 160 g 101 t

ORIGIN

alignment_scores:
Quality: 609.00 Length: 111
Ratio: 5.486 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-13 x BG654378

Align seg 1/1 to: BG654378 from: 1 to: 535

```

1 MetAlaAsnIleHisGlnGluAsnGluGluMetGluGlnProMetGlnAs 17
|||||
177 ATGGCAAAATATTCACCAAGAAACGAAGAGATGGAGCAGCCTATGCAGAA 226
|||||

```

```

17 nGlyGluGluAsnArgProLeuGlyGlyGlyGlyGlyGlyHisGlnProAlaG 34
|||||
227 TGGAGAGGAAGACCGCCCTTTGGAGAGAGTGGAAGCCACACGCTGCAG 276
|||||

```

```

34 LyAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTTPAla 50
|||||
277 GAAATCGACGGGACAGAGCTGCGGACTTGCCCTTAATTTTCGATGGGCC 326
|||||

```

```

51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspAspMe 67
|||||
327 ATACCAATAGGCAGATCAATGATGGATGGGTGGAGATGGAGATGATAT 376
|||||

```

```

67 tGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGluL 84
|||||
377 GGAATATATTCATGGAGAGATGAGAGAAATCAGAAGAAACTTAGGGAGC 426
|||||

```

```

84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
|||||
427 TGCAGTTGAGGAATGTCTGCGTATCTTATGGGGAGCTCTCTAATCAC 476
|||||

```

```

101 HisAspHisHisAspGluPheCysLeuMetPro 111
|||||
477 CATGACCATCATCATCAATTTTGCCTTATGCCT 509
|||||

```

DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 image.llnl.gov/image/html/resources.shtml

Seq primer: -40UP from Gibco
 High quality sequence stop: 422.

FEATURES

Location/Qualifiers
 1..612
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2849369"
 /clone_lib="NCI-CGAP_Lym12"
 /tissue_type="lymphoma, follicular mixed small and large
 cell"
 /lab_host="DH10B"
 /note="Organ: lymph node; Vector: pCMV-SPORT6; Site.1:
 SalI; Site.2: NotI; Cloned unidirectionally. Primer:
 Oligo dt. Average insert size 1.25 kb. Life Technologies
 catalog #: 11547-015"
 catalog # 11547-015"
 BASE COUNT 169 a 145 c 120 g 178 t
 ORIGIN

alignment_scores:
 Quality: 609.00 Length: 111
 Ratio: 5.486 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-13 x AW512400/rev ..

Align seg 1/1 to reverse of: AW512400 from: 1 to: 612

```

1 MetAlaAsnIleHisGlnGluAsnGluMetClnProMetGlnAs 17
|||||
581 ATGCCAAATATTCCACGAGAAACGAGAGATGGAGCGCTATGCAGAA 532

17 nGlyGluAspArgProLeuGlyGlyGlyGlyHisGlnProAla 34
|||||
531 TGGAGAGAGACCGCCCTTTGGGAGAGGTGAAGCCACCGCTGCAG 482

34 lyAsnArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAla 50
|||||
481 GAAATCGACGGGACAGCGCTCGCGACTTGCCTTAATTTTCGATGGGCC 432

51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspPme 67
|||||
431 ATACCCCAATAGGCAGATCAATGATGGGATGGGTGGAGATGGAGATGAT 382

67 tGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGlu 84
|||||
381 GGAATATTTCATGGAGAGATGAGAGAAATCAGAGAAACTTAGGGAGC 332

84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
|||||
331 TGCAGTTGAGGAATTTGTCGCTATCCTTTATGGGGAGCTCTCTAATCAC 282

101 HisAspHisHisAspGluPheCysLeuMetPro 111
|||||
281 CATGACCATCATGATGAATTTTGCCTTATGCCT 249
```

seq_name: gb_est1:A1193112

seq_documentation_block:
 LOCUS A1193112 615 bp mRNA EST 29-OCT-1998
 DEFINITION qe69g09.x1 Soares_fetal_lung_NBHL19W Homo sapiens cDNA clone
 IMAGE:1744288 3' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
 KD PROTEIN HGR74 (HUMAN);, mRNA sequence.

ACCESSION A1193112
 VERSION A1193112.1 GI:3744321
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE AUTHORS TITLE JOURNAL COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 615)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 845 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 445.

FEATURES

source

Location/Qualifiers
 1..615
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1744288"
 /clone_lib="Soares_fetal_lung_NBHL19W"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: lung; Vector: pT73D (Pharmacia) with a
 modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer
 [5'-TGTTACCAATCTGAAGTGGAGCGGCCGCAATTTTTTTTTTTT-3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT73 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M.Fatima Bonaldo. This library was constructed
 from the same fetus as the fetal heart library, Soares
 fetal heart NBHL19W."
 BASE COUNT 169 a 145 c 121 g 180 t
 ORIGIN

alignment_scores:

Quality: 609.00 Length: 111
 Ratio: 5.486 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-13 x A1193112/rev ..

Align seg 1/1 to reverse of: A1193112 from: 1 to: 615

```

1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
|||||
584 ATGCCAAATATTCCACGAGAAACGAGAGATGGAGCGCTATGCAGAA 535

17 nGlyGluClnAspArgProLeuGlyGlyGlyGlyHisGlnProAla 34
|||||
534 TGGAGAGAGACCGCCCTTTGGGAGGAGGTCAAGCCACCGCTGCAG 485

34 lyAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAla 50
|||||
484 GAAATCGACGGGACAGCGCTCGCGACTTGCCTTAATTTTCGATGGGCC 435

51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspPme 67
|||||
434 ATACCCCAATAGGCAGATCAATGATGGGATGGGTGGAGATGGAGATGAT 385

67 tGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGlu 84
|||||
384 GGAATATTTCATGGAGAGATGAGAGAAATCAGAGAAACTTAGGGAGC 335

84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
|||||
334 TGCAGTTGAGGAATTTGTCGCTATCCTTTATGGGGAGCTCTCTAATCAC 285

101 HisAspHisHisAspGluPheCysLeuMetPro 111
```

```

|||||
284 CATGACCATCATGATGAATTTTGGCTTATGCCT 252
seq_name: gb_est1:AA576958

seq_documentation_block:
LOCUS      AA576958      620 bp      mRNA      12-SEP-1997
DEFINITION nm82d04.s1 NCI_CGAP_Co9 Homo sapiens cDNA clone IMAGE:1074727 3'
            similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74
            (HUMAN);, mRNA sequence.
ACCESSION  AA576958
VERSION    AA576958.1 GI:2354432
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 620)
REFERENCE  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.,
            , Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/dbp/image/image.html
            Insert Length: 725 Std Error: 0.00
            Seq primer: -40m13 fwd ET from Amersham
            High quality sequence stop: 339.

FEATURES
            Location/Qualifiers
                1..620
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:1074727"
                /clone_lib="NCI_CGAP_Co9"
                /tissue_type="colon tumor RER+"
                /lab_host="DH10B"
                /note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
                modified polylinker; 1st strand cDNA was prepared from
                RER+ colon tumor, and was then primed with a Not I -
                oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                RI adaptors (Pharmacia), digested with Not I and cloned
                into the Not I and Eco RI sites of the modified pT7T3
                vector. Library is not normalized. Library was
                constructed by Bento Soares and M. Fatima Bonaldo (Soares4
                )."
                169 a 144 c 116 g 191 t

BASE COUNT  169 a 144 c 116 g 191 t
ORIGIN

alignment_scores:
            Quality: 609.00      Length: 111
            Ratio: 5.486        Gaps: 0
            Percent similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-327-750d-13 x AA576958/rev ..
Align seg 1/1 to reverse of: AA576958 from: 1 to: 620

1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
|||||
602 ATGGCAATATTACCAGGAACAGAGAGATGGAGCAGCTATGCAGAA 553
|||||
17 nGlyGluGluAspArgProLeuGlyGlyGlyHisGlnProAlaG 34
|||||
552 TGGAGAGGAAGACCGCCCTTTGGGAGAGGTGAAGCCACCGCCTGCAG 503
|||||
34 LyAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrpAla 50
|||||
502 GAAATCGACGGGACAGCGCTGCCGACTTGCCCTTAATTTTCGATGGGCC 453
|||||
51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyArgGlyAspGly 67
|||||
452 ATACCAATAGGCAGATCAATGATGGTGGTGGTGGATGATGATAT 403
|||||
67 tGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGlu 84
|||||
402 GGAATATTTCATGAGGAGATGAGAGAAATCAGAGAAACCTTAGGGAGC 353
|||||
84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
|||||
352 TGCAGTTGAGGAATGTCTCGTATCCTTATGGGGAGCTCTCTAATCAC 303
|||||
101 HisAspHisHisAspGluPheCysLeuMetPro 111
|||||
302 CATGACCATCATGATGAATTTTGCCTTATGCCT 270
|||||
seq_name: gb_est1:AI929106

seq_documentation_block:
LOCUS      AI929106      662 bp      mRNA      23-AUG-1999
DEFINITION au65b10.y1 Schneider fetal brain 00004 Homo sapiens cDNA clone
            IMAGE:2519611 5' similar to gb:M38188 OVARIAN GRANULOSA CELL 13.0
            KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION  AI929106
VERSION    AI929106.1 GI:5665070
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 662)
REFERENCE  Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
            WashU-NCI human EST Project
            Unpublished (1997)
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40RP from Gibco
            High quality sequence stop: 437.

FEATURES
            Location/Qualifiers
                1..662
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:2519611"
                /clone_lib="Schneider fetal brain 00004"
                /sex="male"
                /tissue_type="frontal lobe"
                /dev_stage="5 months post-conception"
                /lab_host="DH10B"
                /note="Organ: brain; Vector: pBluescript SK (Stratagene);
                Site_1: SstI; Site_2: XhoI; Double-stranded cDNA was
                prepared from human fetal brain tissue. 5' and 3'
                adaptors were used in cloning as follows: 5' adaptor
                sequence:
                5'-GAGAGAGAGAGAGAGCTCAAGGATCCTTAATTAATTAATCCCCCCCCCCC-3'
                and 3' adaptor sequence:
                5'-GAGAGAGAGAGAGCTCGAGTCTTTTTTTTTTTTTTTT-3'. The library was
                size-selected for >0.5 kb inserts and has an average
                insert size estimated at 1.2 kb. This library was
                constructed using the CAP-trapper method for full-length

```


enrichment and has not undergone amplification. Library
was constructed by Dr. Claudio Schneider (LNCIB-Area
Science Park, Trieste, Italy). "

BASE COUNT 178 a 143 c 169 g 170 t 2 others
ORIGIN

alignment_scores:
Quality: 609.00 Length: 111
Ratio: 5.486 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-13 x AI929106 ..

Align seg 1/1 to: AI929106 from: 1 to: 662

```

1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
|||||
96 ATGCCAAATATTCCACGAGGAAACGAAGAGATGGAGCAGCCTATGCAGAA 145

17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyHisGlnProAlaG 34
|||||
146 TGGAGAGGAGAGCCGCCCTTTGGGAGGAGGTGAAGGCCACCCAGCCTGCAG 195

34 lYAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTTPAla 50
|||||
196 GAAATCGACGGGACAGGCTCGCGACTTGCCTTAATTTTCGATGGGCC 245

51 lIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspPme 67
|||||
246 ATACCCAAATAGGCAGATCAATGATGGATGGGTGGAGATGCAGATGATAT 295

67 tGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGluL 84
|||||
296 GGAATATTTCATGGAGGAGATGAGAGAAATCAGAAGAAATTTAGGAGC 345

84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
|||||
346 TGCAGTTGAGGATTTCTGCGTATCTTATGGGGAGGCTCTCTAATCAC 395

101 HisAspHisHisAspGluPheCysLeuMetPro 111
|||||
396 CATGACCATCATGATGAATTTTGCCTTATGCCT 428

```

seq_name: gb_est2:BG745809

seq_documentation_block:

LOCUS BG745809 674 bp mRNA EST 15-MAY-2001
DEFINITION 602723951r1 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:4850293 3',
mRNA sequence.

ACCESSION BG745809

VERSION BG745809.1 GI:14056462

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 674)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Mark Watson

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LCM1691 row: e column: 14

High quality sequence stop: 674.

FEATURES
Source

Location/Qualifiers
1..674

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4850293"

/lab_host="NIH_MGC_113"

/note="Organ: spleen; Vector: pOTB7; Site:1: XhoI; Site:2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the

Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 167 a 186 c 142 g 179 t
ORIGIN

alignment_scores:

Quality: 609.00 Length: 111
Ratio: 5.486 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-13 x BG745809/rev ..

Align seg 1/1 to reverse of: BG745809 from: 1 to: 674

```

1 MetAlaAsnIleHisGlnGluAsnGluGluMetGluGlnProMetGlnAs 17
|||||
529 ATGCCAAATATTCCACGAGGAAACGAAGAGATGGAGCAGCCTATGCAGAA 480

17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyHisGlnProAlaG 34
|||||
479 TGGAGAGGAGAGCCGCCCTTTGGGAGGAGGTGAAGGCCACCCAGCCTGCAG 430

34 lYAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTTPAla 50
|||||
429 GAAATCGACGGGACAGGCTCGCGACTTGCCTTAATTTTCGATGGGCC 380

51 lIleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspPme 67
|||||
379 ATACCCAAATAGGCAGATCAATGATGGATGGGTGGAGATGCAGATGATAT 330

67 tGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGluL 84
|||||
329 GGAATATTTCATGGAGGAGATGAGAGAAATCAGAAGAAATTTAGGAGC 280

84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
|||||
279 TGCAGTTGAGGATTTCTGCGTATCTTATGGGGAGGCTCTCTAATCAC 230

101 HisAspHisHisAspGluPheCysLeuMetPro 111
|||||
229 CATGACCATCATGATGAATTTTGCCTTATGCCT 197

```

seq_name: gb_est1:AV701506

seq_documentation_block:

LOCUS AV701506 720 bp mRNA EST 08-OCT-2000
DEFINITION AV701506 ADB Homo sapiens cDNA clone ADBBBH2 5', mRNA sequence.

ACCESSION AV701506

VERSION AV701506.1 GI:10717836

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 720)

Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao

, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,

Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu


```

84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
|||||
336 TGCAGTTGAGGAATGTCGCGTATCCTTATGGGGAGGCTCTCTAATCAC 287
|||||
101 HisAspHisHisAspGluPheCysLeuMetPro 111
|||||
286 CATGACCATCATGATGAATTTGGCCTTATGCCT 254
|||||
seq_name: gb_est2:BG482054

seq_documentation_block:
LOCUS BG482054 557 bp mRNA EST 21-MAR-2001
DEFINITION 602527934F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4651353 5',
mRNA sequence.
ACCESSION BG482054
VERSION BG482054.1 GI:13414333
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 557)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1434 row: d column: 10
High quality sequence stop: 556.
Location/Qualifiers
1..557
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4651353"
/clone_lib="NIH_MGC_21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: placenta; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT 157 a 126 c 166 g 108 t
ORIGIN
```

```

alignment_scores:
Quality: 597.00 Length: 111
Ratio: 5.427 Gaps: 0
Percent Similarity: 99.099 Percent Identity: 99.099
```

```

alignment_block:
US-09-327-750D-13 x BG482054 ..
```

```
Align seg 1/1 to: BG482054 from: 1 to: 557
```

```

1 MetaAlaAsnIleHisGlnGlnAsnGluMetGluGlnProMetGlnAs 17
|||||
185 ATGCCAATATTCACCAAGAAACAGAGATGGAGCAGCCTATGCAGAA 234
|||||
```

```

17 nGlyGluGluAspArgProLeuGlyGlyGlyHisGlnProAlaG 34
|||||
235 TGGAGAGGAAGACCGCCCTTTGGGAGGAGGTGAAGGCCACCGCCTGCAG 284
|||||
```

```

34 lyAsnArgArgGlyGlnAlaAargLeuAlaProAsnPheArgTtpAla 50
|||||
285 GAAATGACGGGACAGGCTCGCCGACTGCCCTTATTTTCGATGGGCC 334
|||||
51 ileProAsnArgGlnIleLeuAsnAspGlyMetGlyGlyAspGlyAspPme 67
|||||
335 ATACCCAATAGGCAGATCAATGATGGGATGGGTGGAGATGGAGATGAT 384
|||||
67 tGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGluL 84
|||||
385 GGAATATTTCATGGAGAGATGAGAGAAATCAGAGAAACTTAGGGAGC 434
|||||
84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
|||||
435 TGCAGTTGAGGAATGTCGCGTATCCTTATGGGGAGGCTCTCTAATCAC 484
|||||
101 HisAspHisHisAspGluPheCysLeuMetPro 111
|||||
485 CATGACCATCATGATGAATTTGGCCTTATGCCT 517
|||||
```

```
seq_name: gb_est2:BG705685
```

```

seq_documentation_block:
LOCUS BG705685 567 bp mRNA EST 07-MAY-2001
DEFINITION 602668910F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4791506 5',
mRNA sequence.
```

```
ACCESSION BG705685
VERSION BG705685.1 GI:13980274
```

```
KEYWORDS EST.
SOURCE human.
```

```
ORGANISM Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE 1 (bases 1 to 567)
```

```
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
```

```
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
JOURNAL Unpublished (1999)
```

```
COMMENT
```

```
Contact: Robert Strausberg, Ph.D.
```

```
Email: cgapbs-remail.nih.gov
```

```
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
```

```
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
```

```
Toshiyuki and Piero Carninci (RIKEN)
```

```
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

```
DNA Sequencing by: Incyte Genomics, Inc.
```

```
Clone distribution: MGC clone distribution information can be
```

```
found through the I.M.A.G.E. Consortium/LLNL at:
```

```
http://image.llnl.gov
```

```
Plate: L1AM10568 row: d column: 03
```

```
High quality sequence stop: 567.
```

```
Location/Qualifiers
```

```
source
```

```
1..567
```

```
/organism="Homo sapiens"
```

```
/db_xref="taxon:9606"
```

```
/clone="IMAGE:4791506"
```

```
/clone_lib="NIH_MGC_96"
```

```
/tissue_type="hypothalamus"
```

```
/lab_host="DH10B"
```

```
/note="Organ: brain; Vector: pBluescriptR (modified
```

```
pBluescript KS+); Site:1: BamHI; Site:2: Sali-XhoI (gtcgag
```



```
size-selected for average insert size 2.3 kb and
```

```
normalized to 500. This is a primary library enriched
```

```
for full-length clones and constructed using the
```

```
Cap-trapper method (Carninci, in preparation). Library
```

```
constructed by M. Brownstein (NHGRI, National
```

```
Institutes of Health). Note: this is a NIH_MGC Library."
```

```
BASE COUNT 160 a 130 c 168 g 109 t
```

```
ORIGIN
```

```
alignment_scores:
```

```
Quality: 597.00
```

```
Length: 111
```

Ratio: 5.427 Gaps: 0
Percent Similarity: 99.099 Percent Identity: 99.099

alignment_block:

US-09-327-750D-13 x BG705685

Align seg 1/1 to: BG705685 from: 1 to: 567

```

1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
|||||
194 ATGCCAAATATTCCACGAGAAACGAAGAGATGGAGCAGCCTATGCAGAA 243
|||||
244 TGGAGAGAGAGACCGCCCTTTGGGAGAGAGTGGAAGCCACAGCCTGCAG 293
|||||
34 lyAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrrpAla 50
|||||
294 GAAATCGACGGGACAGCGCTTGGGAGAGAGTGGAAGCCACAGCCTGCAG 343
|||||
51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspPme 67
|||||
344 ATACCCAATAGGCAGATCAATGATGGGATGGGTCGAGATGGAGATGATAT 393
|||||
67 tGluIlePheMetGluMetArgGluIleArgArgLysLeuArgGluL 84
|||||
394 GGAATATTTCATGGAGAGATGAGAGAAATCAGAAGAAACTTAGGAGGC 443
|||||
84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
|||||
444 TGCAGTTGAGGAATGTCTCGTATCCTTATGGGGAGGCTCTCTAATCAC 493
|||||
101 HisAspHisAspGluPheCysLeuMetPro 111
|||||
494 CATGACCATCATGATGAATTTGGCCTTATGCCT 526
|||||

```

seq_name: gb_est2:BE896163

```

seq_documentation_block:
LOCUS BE896163 676 bp mRNA EST 20-OCT-2000
DEFINITION 601438756F1 NIH_MGC-72 Homo sapiens cDNA clone IMAGE:3923941 5',
mRNA sequence.
ACCESSION BE896163
VERSION BE896163.1 GI:10360109
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 676)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9760 row: 0 column: 14
High quality sequence stop: 671.
Location/Qualifiers
1. .676
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3923941"
/clone_lib="NIH_MGC-72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;

```

FEATURES

source

```

1. .676
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3923941"
/clone_lib="NIH_MGC-72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;

```

Site_2: Sall; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 179 a 147 c 175 g 175 t
ORIGIN

alignment_scores:

Quality: 597.00 Length: 111
Ratio: 5.427 Gaps: 0
Percent Similarity: 99.099 Percent Identity: 99.099

alignment_block:

US-09-327-750D-13 x BE896163

Align seg 1/1 to: BE896163 from: 1 to: 676

```

1 MetAlaAsnIleHisGlnGluAsnGluMetGluGlnProMetGlnAs 17
|||||
107 ATGCCAAATATTCCACGAGAAACGAAGAGATGGAGCAGCCTATGCAGAA 156
|||||
17 nGlyGluAspArgProLeuGlyGlyGlyGlyGlyHisGlnProAlaG 34
|||||
157 TGGAGAGAGAGACCGCCCTTTGGGAGAGAGTGGAAGCCACAGCCTGCAG 206
|||||
34 lyAsnArgArgGlyGlnAlaArgArgLeuAlaProAsnPheArgTrrpAla 50
|||||
207 GAAATCGACGGGACAGGCTGCGCCCTAATTTTCGATGGGCC 256
|||||
51 IleProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspPme 67
|||||
257 ATACCCAATAGGCAGATCAATGATGGGATGGGTCGAGATGGAGATGATAT 306
|||||
67 tGluIlePheMetGluMetArgGluIleArgArgLysLeuArgGluL 84
|||||
307 GGAATATTTCATGGAGAGATGAGAGAAATCAGAAGAAACTTAGGAGGC 356
|||||
84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
|||||
357 TGCAGTTGAGGAATGTCTCGTATCCTTATGGGGAGGCTCTCTAATCAC 406
|||||
101 HisAspHisAspGluPheCysLeuMetPro 111
|||||
407 CATGACCATCATGATGAATTTGGCCTTATGCCT 439
|||||

```

seq_name: gb_est2:BG993540

```

seq_documentation_block:
LOCUS BG993540 550 bp mRNA EST 13-JUN-2001
DEFINITION PM9-HT0911-080201-012-h10 HT0911 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG993540
VERSION BG993540.1 GI:14397610
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 550)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Negai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

```

TITLE

JOURNAL

MEDLINE

COMMENT

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM0&t2=PM0-HT0911-080201-012-h10&t3=2001-02-08&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 550.

FEATURES

source
 1..550
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="HT0911"
 /dev_stage="Adult"
 /note="Organ: head_neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 154 a 113 c 137 g 146 t
 ORIGIN

alignment_scores:
 Quality: 596.00 Length: 112
 Ratio: 5.369 Gaps: 1
 Percent Similarity: 99.107 Percent Identity: 99.107

alignment_block:

US-09-327-750D-13 x BG993540 ..

Align seg 1/1 to: BG993540 from: 1 to: 550

1 MetAlaAsnIleHisGlnGluAsnGluMetGluGluProMetGlnAs 17
 |||||
 48 ATGGCAAAATATTCCACGAGAAACGAGAGATGGAGCAGCCTATGCAGAA 97
 17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyGlyHisGlnProAlaG 34
 |||||
 98 TGGAGAGAAAGACCGCCCTTTGGGAGGAGGTGAAGGCCACGAGCCTGCAG 147
 34 lyAsnArgArgGlyGlnAlaArgArgLeuAla.ProAsnPheArgTrpAl 50
 |||||
 148 GAAATCGACGGGACAGGCTCGCGGACTTGCCTTAATTTTCGATGGGC 197
 50 aileProAsnArgGlnIleAsnAspGlyMetGlyGlyAspGlyAspAspM 67
 |||||
 198 CATACCAATAGGCAGATCAATGATGGATGGGATGGAGATGGAGATGATA 247
 67 etGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGlu 83
 |||||
 248 TGGAAATATTTCATGGAGGAGATGAGAAATCAGAAAGAAACTTAGGGAG 297
 84 LeuGlnLeuArgAsnCysLeuArgIleLeuMetClyGluLeuSerAsnHi 100
 |||||
 298 CTGCAGTTGAGGAATTCCTCGCTATCCTTATGGGGAGGCTCTCTAATCA 347
 100 sHisAspHisHisAspGluPheCysLeuMetPro 111
 |||||
 348 CCATGACCATCATGATGAATTTTGCCTTATGCCT 381

seq_name: gb_est1:BE252272

seq_documentation_block:

LOCUS BE252272 559 bp mRNA EST 13-JUL-2000
 DEFINITION 60114115F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3354819 5',
 mRNA sequence.
 ACCESSION BE252272
 VERSION BE252272.1 GI:9122409
 KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (Bases 1 to 559)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCM157 row: n column: 04
 High quality sequence stop: 559.

FEATURES

source

1..559
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3354819"
 /clone_lib="NIH_MGC_16"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCAGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."
 BASE COUNT 157 a 126 c 166 g 110 t
 ORIGIN

alignment_scores:
 Quality: 591.00 Length: 111
 Ratio: 5.422 Gaps: 0
 Percent Similarity: 98.198 Percent Identity: 98.198

alignment_block:

US-09-327-750D-13 x BE252272 ..

Align seg 1/1 to: BE252272 from: 1 to: 559

1 MetAlaAsnIleHisGlnGluAsnGluMetGluGluProMetGlnAs 17
 |||||
 177 ATGGCAAAATATTCCACGAGAAACGAGAGATGGAGCAGCCTATGCAGAA 226
 17 nGlyGluGluAspArgProLeuGlyGlyGlyGlyGlyHisGlnProAlaG 34
 |||||
 227 TGGAGAGAAAGACCGCCCTTTGGGAGGAGGTGAAGGCCACGAGCCTGCAG 276
 34 lyAsnArgArgGlyGlnAlaArgArgLeuAla.ProAsnPheArgTrpAla 50
 |||||
 277 GAAATCGACGGGACAGGCTCGCGGACTTGCCTTAATTTTCGATGGGCC 326
 51 ileProAsnArgGlnIleAsnAspGlyMetGlyGlyGlyAspGlyAspPme 67
 |||||
 327 ATACCAATAGGCAGATCAATGATGGATGGGATGGGATGGAGATGATAT 376
 67 tGluIlePheMetGluGluMetArgGluIleArgArgLysLeuArgGluL 84
 |||||
 377 GAAATATTTCATGGAGGAGATGAGAAATCAGAAAGAAACTTAGGGAGC 426
 84 euGlnLeuArgAsnCysLeuArgIleLeuMetGlyGluLeuSerAsnHis 100
 |||||
 427 TGCAGTTGAGGAATTCCTCGTATCCTTATGGGGAGGAGCTCTCTAATCAC 476
 101 HisAspHisHisAspGluPheCysLeuMetPro 111

```
|||||CATGACCATCATGATGAATTTGGCTTATGCCT 509
seq_name: gb_est1:AI928835
seq_documentation_block: 626 bp mRNA EST 23-AUG-1999
LOCUS AI928835
DEFINITION au65b10.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone
IMAGE:2519611.3' similar to gb:M38188 OVARIAN GRANULOSA CELL 13..0
KD PROTEIN HGR74 (HUMAN);, mRNA sequence.
ACCESSION AI928835
VERSION AI928835.1 GI:5664828
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 626)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,X., Wylie,T., Waterston,R. and Willson,R.
WASHU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 475.
FEATURES
Location/Qualifiers
1..626
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2519611"
/clone_lib="Schneider fetal brain 00004"
/sex="male"
/tissue_type="frontal lobe"
/dev_stage="5 months post-conception"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript SK (Stratagene);
Site1: SstI; Site2: XhoI; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence:
5'-GAGAGAGAGAGAGCTCAAGATCCTTAATTAATTAATCCCCCCCCCCC-3'
and 3' adaptor sequence:
5'-GAGAGAGAGAGCTCGAGTTTCTTTTCTTTTCTTTT-3'. The library was
size-selected for >0.5 kb inserts and has an average
insert size estimated at 1.2 kb. This library was
constructed using the CAP-trapper method for full-length
enrichment and has not undergone amplification. Library
was constructed by Dr. Claudio Schneider (LNCIB-Area
Science Park, Trieste, Italy)."
```

BASE COUNT 175 a 144 c 122 g 184 t 1 others

alignment_scores: Quality: 589.00 Length: 111
Ratio: 5.355 Gaps: 0
Percent Similarity: 99.099 Percent Identity: 97.297

alignment_block:
US-09-327-750D-13 x AI928835/rev ..

Align seg 1/1 to reverse of: AI928835 from: 1 to: 626

1 MetAlaAsnIleHisGlnGluAsnGluGluMetGluGlnProMetGlnAs 17

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 14:17:05 ; Search time 1319.57 Seconds
(without alignments)
8751.355 Million cell updates/sec

Title: US-09-327-750D-28

Perfect score: 700

Sequence: 1 acagcgctgtgccagcagc.....ggggctgtgtgccagtga 700

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.om.*

20: em.or.*

21: em.ov.*

22: em.pat.*

23: em.ph.*

24: em.pl.*

25: em.ro.*

26: em.sts.*

27: em.sy.*

28: em.un.*

29: em.vi.*

30: em.htgo_hum.*

31: em.htgo_inv.*

32: em.htgo_rod.*

33: em.htg_hum.*

34: em.htg_inv.*

35: em.htg_rod.*

36: em.htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	692	98.9	700	10	AF187066	AF187066 Mus muscu
2	676	96.6	854	10	AF097440	AF097440 Mus muscu
3	355.8	50.8	519	10	AF187065	AF187065 Rattus no
4	337.8	48.3	793	9	BC003190	BC003190 Homo sapi
5	307.4	43.9	891	9	HU00G	M38188 Human unkno
6	307.4	43.9	891	9	AF187064	AF187064 Homo sapi
7	240.6	34.4	421	11	G72708	G72708 MARC 4953-4
8	154.8	22.1	785	10	AF097439	AF097439 Mus muscu
9	154.6	22.1	98274	9	HS71487	Z99755 Human DNA s
10	101.4	14.5	477	11	G35294	G35294 human STS S
11	101.4	14.5	45678	9	HSV351F8	Z70719 Human DNA s
12	80.2	11.5	1229	9	AK000959	AK000959 Homo sapi
13	80.2	11.5	1364	6	AX100231	AX100231 Sequence
14	80.2	11.5	69648	9	HS635G19	AL035494 Human DNA
15	77.4	11.1	244	11	GL3550	GL3550 human STS S
16	67.6	9.7	99	10	MUSCEREBL	LI0905 Mus musculu
17	59.8	8.5	7218	6	I66494	I66494 Sequence 14
18	58	8.3	642	9	AF183416	AF183416 Homo sapi
19	58	8.3	828	9	AF220189	AF220189 Homo sapi
20	58	8.3	35714	9	HS198P4	AL008708 Human DNA
21	57.2	8.2	43952	9	HS0105G4	Z92846 Human DNA s
22	56.4	8.1	791	9	AF237783	AF237783 Homo sapi
23	55.8	8.0	744	9	AF251053	AF251053 Homo sapi
24	55.8	8.0	898	6	AX078272	AX078272 Sequence
25	55.8	8.0	31321	9	HSV870H8	Z70233 Human DNA s
26	55.8	8.0	41029	2	HS080B1	AL022169 Homo sapi
27	54.4	7.8	504	11	G24641	G24641 human STS W
28	54.2	7.7	40584	9	AL133348	AL133348 Human DNA
29	54	7.7	835	10	AF097438	AF097438 Mus muscu
30	54	7.7	838	10	AF051347	AF051347 Mus muscu
31	54	7.7	2269	10	AF097437	AF097437 Mus muscu
32	49.2	7.0	185306	2	AC013570	AC013570 Homo sapi
33	48.8	7.0	37224	3	AC007061	AC007061 Leishmani
34	48.2	6.9	36823	9	HS0209G1	Z68873 Human DNA s
35	46	6.6	154577	2	AL357046	AL357046 Homo sapi
36	46	6.6	154959	2	AC009282	AC009282 Homo sapi
37	46	6.6	185257	9	AL355305	AL355305 Human DNA
38	44.8	6.4	190544	9	AC005768	AC005768 Homo sapi
39	44.2	6.3	110000	2	LMFLCHR15_5	Continuation (6 of
40	44.2	6.3	172200	2	AC022626	AC022626 Homo sapi
41	43.2	6.2	9211	3	AB030033	AB030033 Dictyoste
42	43.2	6.2	210321	2	AC079438	AC079438 Mus muscu
43	42.8	6.1	1580	3	PFACSPH	J02695 Plasmodium
44	42.6	6.1	164025	2	AC022562	AC022562 Homo sapi
45	42.6	5.1	166922	2	AL512288	AL512288 Homo sapi

ALIGNMENTS

RESULT 1

AF187066

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

AF187066 Mus musculus p75NTR-associated cell death executor (Nade) mRNA, 11-JUN-2000

Complete cds.

AF187066

AF187066.1 GI:8452897

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 700)

Mukai,J., Hachiya,T., Shoji-Hoshino,S., Kimura,M.T., Nadano,D.,

Suvenito,P., Hanaoka,T., Li,Y., Irie,S., Greene,L.A. and Sato,T.A.

NADP, a p75NTR-associated cell death executor, is involved in

signal transduction mediated by the common neurotrophin receptor

p75NTR

J. Biol. Chem. 275 (23), 17566-17570 (2000)

20298829

REFERENCE 2 (bases 1 to 700)
AUTHORS Mukai,J., Hachiya,T., Hoshino,S., Kimura,M., Nadano,D., Suvanto,P.,
Hanaoka,T., Li,Y., Irie,S. and Sato,T.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia
University, 630 West 168th St., P&S II-451, New York, NY 10032, USA

FEATURES
source location/Qualifiers
1..700
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
1..700
/gene="Nade"
/177..551
/gene="Nade"
CDS p75NTR-mediated signal transduction; NADe"
/notes="involved in the common neurotrophin receptor
p75NTR-mediated signal transduction; NADe"
/codon_start=1
/product="p75NTR-associated cell death executor"
/protein_id="AAF75131.1"
/db_xref="GI:8452898"
/translation="MANVHOENEEMEQPLONGOEDRPVGEGGCHOPAANNNNHHNH
HNHRGQARRLAPNRWALPNRMNDGLGGDDMMPEMREIRRKLRLEQLRNC
LRIMGLSNHHHDHDFELCLMP"

BASE COUNT	178 a	187 c	203 g	132 t
ORIGIN				

```

Query Match      98.98;   Score 692;   DB 10;   Length 700;
Best Local Similarity 99.3%;   Pred. No. 3.3e-171;
Matches 695; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY    1  acgagcgtcttgccagcagctcggaagtcctctgcgcgcggcgtgagcggcgccg 60
      |||||
DB    1  ACGAGCGTCGTGCCAGCAGCTCGGAGTCTCTTCGCCGCGCGGGCTGGCAGCGGGCCG 60

QY    61 aggcgcgggacagattgactggaagccgagagtccaggcgcggcggaaattgacagg 120
      |||||
DB    61 AGCGCAGCGGCACAGATTGACTGCAAGCCGAGAGTCCAGCGCGCGGAATTTGACAGG 120

QY    121 aggaactgcgcgaaggataggcccagaatatagcaaccaggaaaataatctcatagg 180
      |||||
DB    121 AGGACTACGGCCGAAGGGATAGGCCCCAGAATAGCAACCAGGAACAATAAAATCTCATATGG 180

QY    181 ccaatgtccacaggaaaaaacgaagagctggagcagccccctgcagaaatggacagaaacc 240
      |||||
DB    181 CCATGTGCCACGAAAACGAGAGATGGAGCAGCCCCCTTGCAATATGGACAGGAAGACC 240

QY    241 gccctgtggaggagggtgagggccaccagcctgtgccaacaacaacaacaacaacacca 300
      |||||
DB    241 GCCCTGTGGGAGGAGGTGAGGGCCACCAGCTGCTGTGCAACAACAACAACAACAACCA 300

QY    301 accataaccacaaccacccgaagagccaggctcgacctgcccttaacttcgat 360
      |||||
DB    301 ACCATAACCACAACACACCGAGAGCGCAGGCTCGCCGACTTGCCCTTAACCTTCGGAT 360

QY    361 ggccattcccaacagcagatgaattcacgggttgtgtgagatggagatgatgaaa 420
      |||||
DB    361 GGCCATTCCCAACAGCGCATGAATACGGGTGTGGGTGGAGATGGAGATGATATGAAAA 420

QY    421 tgttcattggaggagatgagagatccggagaagcttagggagctacagctgagaatt 480
      |||||
DB    421 TGTTTCATGGAGGAGATCAGAGAGATCCGGAGAAAGCTTAGGGAGCTACACCTGAGAAATT 480

QY    481 gtctacgcatcttatggggagctgtctaaccacacgcatcaccatgataatttcgcc 540
      |||||
DB    481 GTCTACGCATCCTTATGGGGAGCTGTCTAACCAACCACGATCACCATGATGAATTTCTGCC 540

QY    541 ttatgccttgacttcgctattcccccctgagatccatactgtgactccgcgtgagccc 600
      |||||
DB    541 TTATGCCTTGACTTCGGGTCAATCCCCTCGAGATCCATACTGTGACTGCCGCTGTAGCCC 600

QY    601 ttccccctgcgattttctcgacatgcctttaatgacccqtttgttggtgaagccctgtgatt 660

```


REMARK
COMMENT

Tissue Procurement: ATCC
CDNA Library Preparation: RubIn Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Co.
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcqsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letlicia Hsiao, Martin Krzywnski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasha van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 6 Row: k Column: 22
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA qi: 7657043.

FEATURES

```
1. 793
/organism="Homo sapiens"
/db_xref="LocusID:27018"
/db_xref="taxon:9606"
/clone="MGC:802 IMAGE:33"
/tissue_type="Eye, retina"
/clone_lib="NIH_MGC_16"
/lab_host="DH10B-R"
/note="vector: pOTB7"
```

CDS

```

/codon_start=1
/codon_start=75NTR-associated cell death executor; ovarian
granulosa cell regulator; 13kD"
/protein_id="AAH0190.1"
/db_xref="GI:13112031"
/translation="NANIHQNEWEQPMONGEDRDPLGGEGGQVAPNRCQARLA
PNFRWAIAPNQINDMGCGDDMEIFMEEMREIRKLRLEQLRNLRLMGELSNHHD
HDFDEFLM"

```

BASE COUNT
ORIGIN

```

Query Match      48.3%; Score 337.8; DB 9; Length 793;
Best Local Similarity 73.8%; Pred. No. 3.7e-78;
Matches 520; Conservative 0; Mismatches 132; Indels 53

```

Qy **Db**

1 acgagctctggccagcagctcgagctcctctgcgcgcgcgaggtctgccagcgggccccg 60
||| |||||||| | ||||| | ||| ||| ||||| ||||| |||||
12 ACTCGGCTCTGGTACCACTCCCGCTGCCTGAGTCGGCGGGCTGCATTCGGCCCCG 71

Oy 61 aggcg-agcgggacagattgactgaagccagaggtccaggcgca-----gcgggaa 112
 || ||||| ||| | | | | | | | | |
Db 72 GGGAAAACGGGAGCAGGTCTCGGAGSCTAAGTGTCTCCGCCGCACCTCGGGCACAA 131

QY 113 ttgacagaggactacgccgaaggatagcccgagatagcaccaggaaacaatatct 172
| | | | | | | | | | | | | | | | | |
Dh 132 TCCGAGCACAAGCAGACATCCAGCATAGCGCCACAAAAACACCACAAAAAATCT 181
| | | | | | | | | | | | | | | | | |

[illegible]

Qy 233 ggaacacgcctgtggaggaggtgaggccaccagcctgtgcaaacacaacacaa 292

Qy 293 caaccacaaccataaccacaaccaccaccaggagggccaggctcgccgacctgccctaa 352

Db	299	-----ATCGACGGGGACAGGCTCGCGACTTGCCCCCTAA	332
Qy	353	cttcgatggccattcccaacagcgagatgacgggtgtgggtgagatgagatga	412
Db	333	TTTTTCGATGGCCATACCAATAGCGAGATCAATGATGGGATGGGTGGAGATGGAGATGA	392
Qy	413	tatgaaatgttcatacgagatgagagatgccggagaaagcttaggagctacagct	472
Db	393	TATGAAATATTATCGAGGAGATGAGAAATCAGAAAGAAACTTAGGAGCTGCAGTT	452
Qy	473	gagaaatgtctaacgatacttattggggggagctgtctaacaccacacatcacatgatga	532
Db	453	GAGAAATGTCTCGTATCTCTATGGGGAGCTCTCTAATCACCATGACCATCATGATGA	512
Qy	533	attgcctcatgcttgactcgggtcatctccccctcgagatccatactatgactcccg	592
Db	513	ATTTTGGCTTATGCGCTTGACTCTGCCATT-TATCATGAGATTAACTGTGATTCGCCG	571
Qy	593	tgtagcctttccctcgcatttctctgacatgcctttaatgacccgcttggtaggagccc	652
Db	572	TGTTTTCTTTTCTTCGATTTCCTAATATGCTTTACTGATCCGTTTCTGCTGGAACCC	631
Qy	653	tgtgtatttcacatgcgatgcccaggcggggctgtgctgtgccag	697
Db	632	TATGTTATTT-----CCATGTGCAAGTGGGCTCTTGTGTGCCAG	671
RESULT	5		
HUMOGC			
LOCUS			
DEFINITION		891 bp mRNA	PRI 07-MAR-1995
ACCESSION		Human unknown protein from clone pHR74	mRNA, complete cds.
VERSION		M38188.X56942	
KEYWORDS		M38188.1 GI:189378	
SOURCE		Human ovarian granulosa cell line, cDNA to mRNA.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi.	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE		1 (bases 1 to 891)	
JOURNAL		Rapp,G., Freudenstein,J., Klaudiny,J., Mucha,J., Wempe,F.,	
MEDLINE		Zimmer,M. and Scheit,K.H.	
COMMENT		Characterization of three abundant mRNAs from human ovarian granulosa cells	
FEATURES		DNA Cell Biol. 9 (7), 479-485 (1990)	
source		91025550	
		Draft entry and computer-readable sequence for IDNA 9, 479-485	
		(1990)] kindly submitted	
		by K.H.Scheit, 27-AUG-1990.	
		Location/Qualifiers	
		1..891	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="pHR74"	
		/cell_type="granulosa"	
		/tissue_type="ovary"	
mRNA		<1..891	
		/note="protein of unknown function"	
		312..647	
CDS		/note="protein of unknown function"	
		/codon_start=1	
		/protein_id="AAA63232.1"	
		/db_xref="GI:189378"	
		/translation="MANIQHEMEFOPMQNEEDRPPLGGGGHQHPAGNRRCQARRLA	
		PNFRWAIPIRQINDGMGGDDMEIFMEERIRKRLRELQRLNCLRLIMLGELSNHHD	
		HHDEFCLMP"	
		361..534	
CDS		/note="protein of unknown function"	
		/codon_start=1	
		/protein_id="AAA63233.1"	
		/db_xref="GI:189380"	
		/translation="MERKTAWEVKYATSLQIEDGDRDLADLPLIFDGPYPPIGRSMG	
		VENEMIWKYSRR"	
BASE COUNT	251 a	182 c	234 t

BASE COUNT

misc_feature /note="MER58A repeat: matches 9. .214 of consensus"
13338. .13833 /note="match: GSS: Em:AQ568864"
repeat_region 13686. .13749
repeat_region /note="MIR repeat: matches 77. .139 of consensus"
13840. .13966 /note="MIR repeat: matches 22. .147 of consensus"
repeat_region 14162. .14245 /note="MIR repeat: matches 46. .136 of consensus"
repeat_region 14925. .15038 /note="MIR repeat: matches 48. .191 of consensus"
repeat_region 15994. .16040 /note="MIR repeat: matches 102. .144 of consensus"
repeat_region 16116. .16184 /note="MER58B repeat: matches 3. .70 of consensus"
repeat_region 16897. .17005 /note="MIR repeat: matches 9. .118 of consensus"
repeat_region 17618. .17700 /note="MIR repeat: matches 108. .192 of consensus"
repeat_region 17748. .17912 /note="MIR repeat: matches 73. .245 of consensus"
repeat_region 18657. .18871 /note="MIR repeat: matches 35. .252 of consensus"
repeat_region 18927. .19230 /note="AluX repeat: matches 1. .301 of consensus"
repeat_region 19492. .19559 /note="MIR repeat: matches 2. .67 of consensus"
repeat_region 19560. .19821 /note="AluY repeat: matches 34. .295 of consensus"
repeat_region 19822. .19959 /note="MIR repeat: matches 67. .212 of consensus"
repeat_region 19987. .20286 /note="AluY repeat: matches 1. .300 of consensus"
repeat_region 20316. .20512 /note="MIR repeat: matches 82. .262 of consensus"
repeat_region 20513. .20741 /note="L1MD1 repeat: matches 5970. .6224 of consensus"
repeat_region 20769. .21074 /note="AluSg1 repeat: matches 1. .303 of consensus"
repeat_region 21085. .21207 /note="L1M2 repeat: matches 1776. .1898 of consensus"
repeat_region 21212. .21391 /note="L1MD repeat: matches -6. .178 of consensus"
repeat_region 21392. .21471 /note="MIR repeat: matches 3. .86 of consensus"
repeat_region 22146. .22441 /note="AluX repeat: matches 1. .296 of consensus"
repeat_region 22526. .22610 /note="LTR16C repeat: matches 267. .349 of consensus"
repeat_region 22940. .23204 /note="HERV16 repeat: matches 378. .650 of consensus"
repeat_region 23205. .23514 /note="AluSg repeat: matches 1. .311 of consensus"
repeat_region 23919. .24194 /note="AluX repeat: matches 36. .311 of consensus"
repeat_region 24496. .24590 /note="LTR16C repeat: matches 253. .349 of consensus"
repeat_region 24871. .24916 /note="MIR repeat: matches 102. .143 of consensus"
repeat_region 25378. .25555 /note="L2 repeat: matches 2330. .2519 of consensus"
repeat_region 25673. .25910 /note="MIR repeat: matches 4. .255 of consensus"
repeat_region 26722. .26798 /note="L2 repeat: matches 2572. .2649 of consensus"
repeat_region 27040. .27107 /note="L2 repeat: matches 2489. .2500 of consensus"
repeat_region 27108. .27374 /note="AluY repeat: matches 41. .304 of consensus"
repeat_region 27375. .27471 /note="L2 repeat: matches 2372. .2489 of consensus"
repeat_region 27651. .27945 /note="AluX repeat: matches 1. .299 of consensus"

repeat_region 28235. .28876 /note="L2 repeat: matches 1113. .1817 of consensus"
repeat_region 28975. .29097 /note="MER96 repeat: matches 42. .173 of consensus"
repeat_region 29173. .29196 /note="L2 copies 2 mer aa 100 conserved"
misc_feature 29832. .30260 /note="match: GSS: Em:AQ075670"
repeat_region 30392. .30597 /note="MIR repeat: matches 35. .252 of consensus"
repeat_region 30845. .31010 /note="MER91B repeat: matches 2. .162 of consensus"
repeat_region 31056. .31094 /note="MIR repeat: matches 218. .256 of consensus"
repeat_region 31728. .31791 /note="L6 copies 4 mer tctg 76 conserved"
repeat_region 31858. .31952 /note="MIR repeat: matches 48. .153 of consensus"
misc_feature complement(33437. .33589) /note="match: STS: Em:H55731"
repeat_region 33543. .33633 /note="MIR repeat: matches 84. .165 of consensus"
repeat_region 33785. .33966 /note="MER91A repeat: matches 1. .185 of consensus"

Query Match 22.1%; Score 154.6; DB 9; Length 98274;
Best Local Similarity 76.5%; Pred. No. 7e-30; Mismatches 0; Indels 14; Gaps 6;
Matches 270; Conservative 0;

QY 345 gccctaaattccgagtgccattcccaacaggcagatgaatgacgggttggtggagat 404
Db 62890 GCCCTAAATTTCAATGGCCATACCAATAGCAGGTCAATGATGGATAGAGGT 62949

QY 405 gaagatgatataatgaaatgttcattgaggagatgagagagatcccgagaagattagggg 464
Db 62950 GGAGATGATATGCAATGTTTCATGGAGGAGATGAGAGGAATCAGGAG--AGAATAAGGAG 63007

QY 465 ctacagctgagaaattgtctacgcattctatggggagctgtctaaaccacgatac 524
Db 63008 CTACAAATGAGGAATTTGCTGTATCCTTAT--GGAAAGCTGTGTAATCCCCATGACCAT 63066

QY 525 catgatgaattcgcttatgccttgacttcgctgattccctccctcgagatccatctgtg 584
Db 63067 CATGATGAATTTTGGC--AATGCTGACTCTCCATTT--TTCCATGAGGTTAATATTGTG 63124

QY 585 atccgcgtgtgaccttccctcgcatttctcctgacatgcctttaaagaccgcttctg 644
Db 63125 ATTCCATTTGTTTCTTTTCTTTCCTTTCATTTCTCAATATGCTCTTT---ACTCATTTGCT 63180

QY 645 gtgagccctgttatttccatgcccattgcccaggtgggcttctgttgcacg 697
Db 63181 GTGAACCTTATGTTATTT-----CCATGTGTCAGGTGAGTCTTTGCTTTCCAG 63228

RESULT 10
G35294/c
LOCUS G35294 477 bp DNA STS 02-OCT-1997
DEFINITION human STS SHGC-37409, sequence tagged site.
ACCESSION G35294
VERSION G35294.1 GI:2459462
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 477)
AUTHORS Myers,R.M.
JOURNAL Unpublished (1997)
COMMENT Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA

Db	256	AGAGGGGACAGCCCCCAGCAGATGAAGAAGATCCGCCATTGGGAGGGGGTGA	315			
Qy	260	ggggccacagcctgtgtgcaacaacaacaacaacaacataaccacaaccacca	319			
Db	316	AGCCAGAAGCCTGGAGGAATATCA-----	341			
Qy	320	ccgaagaagccagggctgcgcgcacctgccttaactccgatgggcaattcccaacaggca	379			
Db	342	--GGCGGGGCGAGTATAGGCGACTGTGCCCTAAATTTTCGATGGGCCATACCTAATAGGCA	399			
Qy	380	gatgaatgacgggttggtggagatgatgatatgaaatgttcatgaggagatgag	439			
Db	400	TATTGACCACAA-----TGAGCGAGAGATGATGTAGAAGGTTGTAGGGCAGATGAT	453			
Qy	440	agagatccgagaagaagcttagggagctcacagctgagaaattgtctacgcaccttatggg	499			
Db	454	GGAATCAACGAAGAAGACTAGGGAACAGCAGATGAGGCACCTATATGCGCTCCAAA----	509			
Qy	500	ggagctgtctaacaccacagatcaccatgatgaattctgcttctgactgacttcgctc	559			
Db	510	-----CTCCTGAACCTGACAACCAATTATGACTTTTGCTCTACCTTGAATCCTTAAA	561			
Qy	560	attccccctgagatccatactgtgactcccgctgtagcccttccctcgcatcttct	618			
Db	562	AGTTTCGCTGAGGTTAATGTGAACACTGCTTTACAAGCTTGATATTTTGTGATTACT	620			
RESULT 13						
AX100231						
LOCUS	AX100231	1364 bp DNA PAT	02-APR-2001			
DEFINITION	Sequence 32 from Patent WO0119860.					
ACCESSION	AX100231					
VERSION	AX100231.1 GI:13539115					
KEYWORDS	human.					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 1364)					
AUTHORS	Tang,Y.T., Hillman,J.L., Yue,H., Reddy,R., Lal,P., Shah,P., Azimzai,Y., Baughn,M.R., Lu,D.A., Bandman,O., Shih,L.L. and Patterson,C.					
TITLE	Proteins associated with cell differentiation					
JOURNAL	Patent: WO 0119860-A 32 22-MAR-2001;					
FEATURES	Incyte Genomics, Inc. (US)					
source	Location/Qualifiers					
	1..1364					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	/notes="Incyte ID No: 1990956CB1"					
BASE COUNT	411 a	269 c	322 g			
ORIGIN	362 t					
Query Match 11.5%; Score 80.2; DB 6; Length 1364;						
Best Local Similarity 56.2%; Pred. No. 1.8e-10;						
Matches 269; Conservative 0; Mismatches 153; Indels 57; Gaps 4;						
Qy	143	gccagaatgcaaccagggaacaaaatctcatatgccaattgccaccaggaaaacga	202			
Db	255	GTCAAAAGAGAACTACGCGCAAACTCTCAACGGGAAAATCCCAACAAGAAAACGA	314			
Qy	203	agagctgagcag--cccctgcagaatgacaggaacacgcctctggagagatga	259			
Db	315	AGNAGGGGACGAGGCCCCACGCAAGATGAAGAAGATCCCGCCATTGGGAGGGGGTGA	374			
Qy	260	gggcccacagcctgtctgcaacaacaacaacataaccataaccacaaccacca	319			
Db	375	AGGCCAGAAGCCTGGAGGAATATCA-----	400			
Qy	320	ccgaagggccaggtcgcgcgcacctgccttaactccgatgggcaattcccaacaggca	379			
Db	401	--GGCGGGGCGAGTATAGGCGACTGTGCCCTAAATTTTCGATGGGCCATACCTAATAGGCA	458			

```
/chromosome="X"
/map="q22.1-22.3"
/clone="RP4-635G19"
/clone_lib="RPCI-4"
439..483
/notes="TAR1 repeat: matches 707..749 of consensus"
501..661
/notes="AluJb repeat: matches 137..311 of consensus"
662..897
/notes="118 copies 2 mer aa 70% conserved"
915..1331
/notes="MSTA repeat: matches 1..419 of consensus"
1334..1517
/notes="92 copies 2 mer ga 77% conserved"
1927..2020
/notes="MIR repeat: matches 109..217 of consensus"
2066..2262
/notes="LIME2 repeat: matches 5825..6022 of consensus"
2263..2574
/notes="AluSp repeat: matches 1..313 of consensus"
2575..2819
/notes="LIME2 repeat: matches 5583..5825 of consensus"
2821..3039
/notes="Alu repeat: matches 1..311 of consensus"
3104..3384
/notes="LIME2 repeat: matches 5304..5591 of consensus"
3391..4007
/notes="L1MB1 repeat: matches 5558..6162 of consensus"
4012..4137
/notes="rigger2a repeat: matches 11..123 of consensus"
4141..4252
/notes="L1 repeat: matches 2920..3031 of consensus"
4265..4350
/notes="HY4 repeat: matches 1..87 of consensus"
4350..4678
/notes="MER74A repeat: matches 28..369 of consensus"
4914..5095
/notes="MER5A repeat: matches 8..189 of consensus"
5224..5555
/notes="L2 repeat: matches 2423..2746 of consensus"
6040..6086
/notes="L2 repeat: matches 1676..1720 of consensus"
7413..7478
/notes="Alu repeat: matches 239..304 of consensus"
7628..7928
/notes="AluYb8 repeat: matches 1..316 of consensus"
7952..7987
/notes="18 copies 2 mer aa 81% conserved"
7990..8305
/notes="AluJo repeat: matches 1..309 of consensus"
8436..8538
/notes="MIR repeat: matches 99..212 of consensus"
10836..11442
/notes="L1MD repeat: matches 1357..1972 of consensus"
11443..11571
/notes="FLAM.A repeat: matches 2..129 of consensus"
11572..11946
/notes="L1MD repeat: matches 977..1357 of consensus"
11947..12245
/notes="AluY repeat: matches 1..301 of consensus"
12246..13098
/notes="L1MD repeat: matches 275..977 of consensus"
13094..13223
/notes="L1MD repeat: matches -17..113 of consensus"
13523..13820
/notes="MLTID repeat: matches 177..505 of consensus"
13823..13932
/notes="55 copies 2 mer ct 77% conserved"
13936..13977
/notes="MLTID repeat: matches 201..241 of consensus"
13978..14240
/notes="AluSg1 repeat: matches 1..270 of consensus"
14245..14534
/notes="AluSc repeat: matches 1..290 of consensus"
14535..14721
/notes="MLTID repeat: matches 2..201 of consensus"
14724..15604
/notes="L1MC/D repeat: matches 4646..5565 of consensus"
15605..15909
/notes="AluX repeat: matches 24..312 of consensus"
16069..16363
/notes="AluX repeat: matches 1..291 of consensus"
16482..16641
/notes="AluJb repeat: matches 136..295 of consensus"
16671..16953
/notes="AluX repeat: matches 1..284 of consensus"
17164..17280
/notes="L1P repeat: matches 4499..4619 of consensus"
17326..17377
/notes="26 copies 2 mer at 79% conserved"
17704..17863
/notes="AluSc repeat: matches 118..295 of consensus"
17864..18175
/notes="AluYb8 repeat: matches 1..318 of consensus"
18176..18298
/notes="AluSc repeat: matches 1..118 of consensus"
18848..18873
/notes="13 copies 2 mer tg 100% conserved"
18876..19171
/notes="AluSg repeat: matches 1..295 of consensus"
19183..19232
/notes="25 copies 2 mer ac 100% conserved"
19246..19533
/notes="AluY repeat: matches 5..292 of consensus"
19539..19837
/notes="MER74B repeat: matches 61..368 of consensus"
20016..20194
/notes="MER74A repeat: matches 295..487 of consensus"
20394..20692
/notes="AluSg repeat: matches 1..302 of consensus"
20701..20809
/notes="LTR1 repeat: matches 682..785 of consensus"
20877..21061
/notes="LTR1 repeat: matches 1..195 of consensus"
21186..21636
/notes="LTR29 repeat: matches 3..498 of consensus"
21609..21737
/notes="MER34 repeat: matches 409..543 of consensus"
21771..22076
/notes="AluY repeat: matches 2..308 of consensus"
22878..22974
/notes="MIR repeat: matches 92..191 of consensus"
23452..23582
/notes="FLAM.C repeat: matches 1..131 of consensus"
23583..23624
/notes="21 copies 2 mer ta 76% conserved"
23735..24041
/notes="AluSg1 repeat: matches 1..308 of consensus"
24045..24072
/notes="14 copies 2 mer tt 89% conserved"
25274..25582
/notes="AluJb repeat: matches 1..289 of consensus"
25583..25747
/notes="match: GSS B63622"
25959..26012
/notes="HERVPH21 repeat: matches 23..79 of consensus"
26144..26456
/notes="AluY repeat: matches 1..311 of consensus"
27281..27432
/notes="AluJb repeat: matches 137..289 of consensus"
27436..27507
/notes="L1M4 repeat: matches 5360..5434 of consensus"
27536..27600
/notes="L1M4 repeat: matches 2220..2283 of consensus"
27601..27975
/notes="MLT1A repeat: matches 3..365 of consensus"
```

repeat_region 27976. .28273

Query Match 11.5%; Score 80.2; DB 9; Length 69648;
Best Local Similarity 56.2%; Pred. No. 2.4e-10;
Matches 269; Conservative 0; Mismatches 153; Indels 57; Gaps 4;
QY 143 gccagaatagcaaccaggaaacaaatctcatctatggtcccaatgtccaccaggaaacga 202
Db 64298 GTCCAAAGAGGAAC TAGCGGCAACAACTCAACGGGAAATGCCCAACAAGAAACGA 64357
QY 203 agagctggagcag---ccctgcagaatggacagaaacacccctgtgtggggagggtga 259
Db 64358 AGGAGGGAGCAGGCGCCGCCACGACGAATGAAGAAGATCCCGCCATTTGGGAGGGGTGA 64417
QY 260 gggccaccagcctgtgcacaacaaacaaacacacacacacacacacacacacaccca 319
Db 64418 AGCCAGAGCCTGGAGGAATATCA----- 64443
QY 320 ccgaagagccaggtcgccgaccccttaactccgatggccattcccaacaggca 379
Db 64444 --GGCGGGGGGAGTTAGGCGACTTGTCCCTAATTTTCGATGGCCATACCTAATAGGCA 64501
QY 380 gatgaatgacgggttgggtggagatggagatgatgaaatgttcattgagagagatgag 439
Db 64502 TATTGAGCACA-----TGAAGCGAGAGATGATGTAGAAAGGTTTGTAGGCGAGATGAT 64555
QY 440 agagatccggagaaagcttagggagctacagctgagaaattgtctacgcacatcctatggg 499
Db 64556 GGAATCAGAGAAAGACTAGGGAACAGCAGATGAGGCACTATATGCGGTTCCAAA----- 64611
QY 500 ggaagtgttaaccaccagatcaccatgatgaatttgccttatgcttgccttgcctcggtc 559
Db 64612 -----CTCCTGAACCTGACAACCATATGACATTTTCCCTCATACCTTGAATCCTAAA 64663
QY 560 attccccctgagatccatactgtgactcccgctgtgaccccttccctgcgatttccct 618
Db 64664 AGTTTTCGTGAGGTTAATGTGAACACTGCTTTTACAAGCTTGTATATTTTGTGATTACT 64722

RESULT 15
LOCUS G13550 244 bp DNA STS 04-JAN-1996
DEFINITION human STS SHGC-11140.
ACCESSION G13550
VERSION G13550.1 GI:1129289
KEYWORDS STS sequence; primer; sequence tagged site.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
Catarrhini; Homnidae; Homo.
1 (bases 1 to 244)
Myers.R.M.
Unpublished (1995)
Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu
Primer A: CCGTTTCCTGTGAACCCCTAT
Primer B: TTTAAATTTTCCATGCAATGG
STS size: 138
PCR Profile:
Initial incubation: 94 degrees C for 90 seconds
Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds

PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600
Protocol:
Template: 25 ng each 1 uM
Primer: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 10 ul
Buffer: MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs derived from M38188 -- Unigene.
FEATURES
Location/Qualifiers
1..244
/organism="Homo sapiens"
STS
primer_bind 84..221
primer_bind 84..103
primer_bind complement(199..221)
BASE COUNT 57 a 50 c 42 g 95 t
ORIGIN

Query Match 11.1%; Score 77.4; DB 11; Length 244;
Best Local Similarity 80.2%; Pred. No. 8.6e-10;
Matches 105; Conservative 0; Mismatches 21; Indels 5; Gaps 1;
QY 567 cctgagatccatactgtgactcccgctgtgaccccttccctgcgatttccctgacatgcc 626
Db 15 CATGAGATTAATACGTGATTCGCCGTGTTTCTTTTCTTCCATTTCCTAATATGCC 74
QY 627 tttaatgacccgtttgtgtgagccctgtgttatttccatgccaatgtgccagtgagggt 686
Db 75 TTTACTGATCCGTTTGTCTGTGAACCCCTATGTTATTT-----CCATGTGTCAAGTGGGTCT 129
QY 687 tgtgttgccag 697
Db 130 TGTGTGCCCAG 140

Search completed: March 11, 2002, 16:33:03
Job time: 8158 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 13:08:25 ; Search time 1609.39 Seconds
(without alignments)
4673.848 Million cell updates/sec

Title: US-09-327-750D-28

Perfect score: 700

Sequence: 1 acagagctgtgccagcagc.....ggggctgtgtgccagtga 700

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	687.4	98.2	809	12 AK004531	AK004531 Mus muscu
2	687.4	98.2	845	12 AK010500	AK010500 Mus muscu
3	637	91.0	787	11 BF608455	BF608455 MYL_00139
4	627.6	89.7	642	10 A1118980	A1118980 ue93c08.y
5	606.8	86.7	628	10 AW476468	AW476468 uq76d02.y
6	591	84.4	599	11 BG088461	BG088461 H3153D07-y
7	590.4	84.3	612	10 AA272375	AA272375 vb62g07.r
8	585.4	83.6	826	12 AK003294	AK003294 Mus muscu
9	578	82.6	817	11 BF407164	BF407164 602919012
10	571.6	81.7	585	10 BE334877	BE334877 us90b10.y
11	566.6	80.9	923	11 BF178306	BF178306 601808545
12	535.6	76.5	542	10 BE334866	BE334866 us90a11.y

13	533.2	76.2	810	11 BF168928	BF168928 601775392
14	533.2	76.2	918	11 BF1249717	BF1249717 602996190
15	531.4	75.9	557	10 AA215070	AA215070 mu76f04.r
16	524.8	75.0	577	10 AA268306	AA268306 va89c10.r
17	523.6	74.8	538	10 A1008575	A1008575 uel5e06.y
18	516	73.7	590	11 W64711	W64711 md71g01.r1
19	510.4	72.9	547	10 AW476461	AW476461 uq76c02.y
20	509.8	72.8	543	10 AA050176	AA050176 mj14b06.r
21	509	72.7	577	11 BF681983	BF681983 602116954
22	507	72.4	536	11 W98871	W98871 mf89d04.r1
23	507	72.4	616	11 W81757	W81757 me95d06.r1
24	497.4	71.1	542	10 AA004191	AA004191 mg82d06.r
25	496.8	71.0	1042	11 W15845	W15845 mb53b09.r1
26	492.4	70.3	503	10 AW908751	AW908751 uf57a05.y
27	489.2	69.9	853	11 BF082462	BF082462 602877692
28	487	69.6	522	10 AA068855	AA068855 mm64d09.r
29	485.4	69.3	506	10 AA049613	AA049613 mj36b06.r
30	485.4	69.3	603	10 AW557617	AW557617 L0284B06-r
31	484	69.1	492	11 W46041	W46041 mc82h02.r1
32	483.6	69.1	533	10 AA606578	AA606578 vm82d02.r
33	480	68.6	532	10 AA214909	AA214909 mu76f05.r
34	479.8	68.5	534	10 AA259353	AA259353 va50g11.r
35	476.6	68.1	533	11 W29675	W29675 mc07d02.r1
36	468.4	66.9	489	10 AA044548	AA044548 mj12c02.r
37	466.6	66.7	560	10 AA870604	AA870604 vq23g08.r
38	465	66.4	486	10 AA259786	AA259786 va87c12.r
39	461.2	65.9	509	10 AA636864	AA636864 vr20g04.r
40	458	65.4	540	10 AA466788	AA466788 vd90b09.r
41	454	64.9	473	11 W84975	W84975 mr42g01.r1
42	453.4	64.8	662	11 BF681719	BF681719 602116954
43	452.2	64.6	459	10 BE654418	BE654418 UI-M-AH1-
44	451	64.4	459	10 BE198421	BE198421 uq78d07.y
45	450.6	64.4	533	10 AA214902	AA214902 mu76d05.r

ALIGNMENTS

RESULT 1

AK004531

LOCUS AK004531 809 bp mRNA

DEFINITION Mus musculus 18 days embryo cDNA, RIKEN full-length enriched library, clone:1190011J23, full insert sequence.

ACCESSION AK004531

VERSION AK004531.1 GI:12835760

KEYWORDS CAP trapper.

SOURCE Mus musculus (strain:C57BL/6J) 18 days embryo cDNA to mRNA, clone:1190011J23.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 809)

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Methods in enzymology. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2 (bases 1 to 809)

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome research. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3 (bases 1 to 809)

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kawanishi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multipicapillary sequencer
 Genome research. 10 (11), 1757-1771 (2000)
 20530913
 11076861

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

4 (bases 1 to 809)
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5 (bases 1 to 809)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M.,
 Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F.,
 Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J.,
 Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T.,
 Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y.,
 Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H.,
 Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
 Tanaka, T., Tejima, Y., Toya, T., Yamamura, Y., Yasunishi, A.,
 Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 Please visit our web site (http://genome.gsc.riken.go.jp/) for
 further details.
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues. First strand cDNA was primed with a primer
 [5', GAGAGAGAGCGCGCGACGCGAGCTTTTCTTTTCTTTT 3'], cDNA was
 prepared by using trehalose thermo-activated reverse transcriptase
 and subsequently enriched for full-length by cap-trapper. cDNA went
 through one round of normalization to Rot = 5.0. Second strand cDNA
 was prepared with the primer adapter of sequence [5'
 GAGAGAGAGAGATCAAGAGCTCAATTAAATTAATTAACCCGCCGCC 3']. cDNA was
 cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end:
 XhoI. Host: SOLR.
 Location/Qualifiers
 1..809
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /db_xref="MGD:MGI:1338016"
 /db_xref="MGD:MGI:1892692"
 /clone="1190011J23"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="18 days embryo"
 181..624
 /note="putative"
 /codon_start=1
 /protein_id="BAB23350.1"
 /db_xref="GI:12835761"
 /translation="MANVHOENEWEOPLONGEDRPVGGEGHOPANNNNNNHHN
 NHRRRGARLPNFRWAIPIRQNDGLGGDDMEFMEMREIRKRLRELQLRNC
 LRLMGLSNLHDDHDFCLPWLKSPFPEIHTVTPAVALFAPS"
 790..795
 /note="putative"
 809
 /note="putative"
 208 a 211 c 227 g 163 t
 polyA_signal
 polyA_site
 BASE COUNT 208 a 211 c 227 g 163 t
 ORIGIN

Query Match 98.2% Score 687.4; DB 12; Length 809;

Best Local Similarity 99.1%; Pred. No. 2.3e-152;
 Matches 691; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 acgagcgtctgcccagcagctcgagagctcctctcgcgcgcgcggtgacgagggccg 60
 |||||
 DB 5 ACGAGCGTCTGGCCAGCAGCTCGGAGCTCTCTGCGCGGGGGTGCAGCGGCCG 64
 |||||

QY 61 agcgagcgggacagattgactggaagccgagagtcacagcgcgaggaattgacag 120
 |||||
 DB 65 AGCGAGCGGGACACATTGACTTGGAAAGCCGACAGTCCAGCGCGGGAATTGACAG 124
 |||||

QY 121 aggactagccccaaggatagcccagaatagcaaccaggagaaacaaatctcatcg 180
 |||||
 DB 125 AGGACTACGCCCAAGGATAGCCAGAAATAGCAACAGGAACAAAATCTCATCGG 184
 |||||

QY 181 ccaatgctccaccagaaagaagagctggagcagccctcgagatgagcaggaacacc 240
 |||||
 DB 185 CCAATGTCCACAGGAAACGAAGAGATGAGCAGCCCTGGCAGATGACAGGAGACC 244
 |||||

QY 241 gccctgtggagagtgagggccaccagcctgctgcaaacacaacaacaacacaca 300
 |||||
 DB 245 GCCCTGTGGAGAGGTGAGGGCCACACCCCTGCTGCAAAACAACAACAACACCA 304
 |||||

QY 301 accataaacacacaccacccaggaagggccaggtctgcgcgaactgcccctaactccgat 360
 |||||
 DB 305 ACCATAAACACACACCACCCAGAGAGGCGCAGGCTCGCCGACTTGCCCTAACTTCGAT 364
 |||||

QY 361 gggccattcccaacagggagatgaatgacgggttgggtggagatggagatgataggaaa 420
 |||||
 DB 365 GGGCCATTTCCCAACAGGCAGATGAATGACGGGTGGGTGGAGATGATATGGA 424
 |||||

QY 421 ttttcagagagatgagagagatcccgagaaagcttaggagctacagctgagaaatt 480
 |||||
 DB 425 TTTTCATGAGAGATGAGAGAGATCCGGAGAAAGCTTAGGAGCTTACAGCTGAGAAAT 484
 |||||

QY 481 gtctacgcatccttatggggagctgtctaacacacacacacacacacacacacacacac 540
 |||||
 DB 485 GTCTACGATCCTTATGGGGAGCTGTCTAACACACACACACACACATCATTCTGCC 544
 |||||

QY 541 ttatccttgactcgtgctatcccccctcgagatcctactgtgactcccgctgtagccc 600
 |||||
 DB 545 TTATGCTTGGCTTCGGTCAATTTCCCGCTGAGATCCATCTGACTCCCGCTGAGCCC 604
 |||||

QY 601 ttctcctgcatttccctgacatgcctttaaagaccgtttgtgtagccctgtgttat 660
 |||||
 DB 605 TTTTCTCGCATTTTCTGACATGCTTTAATGACCCGTTTGTGGTGACCCCTTGTGTTAT 664
 |||||

QY 661 ttccatgcatgcccaggtgggggttgggtgtgtgtgccag 697
 |||||
 DB 665 TTCCATGCATGTGCCAGGTGGGGCTTGTGTGGCCAG 701
 |||||

RESULT 2
 AK010500
 LOCUS AK010500 845 bp mRNA HTC 05-JUL-2001
 DEFINITION Mus musculus ES cells cDNA, RIKEN full-length enriched library,
 clone:2410015K23, full insert sequence.
 ACCESSION AK010500
 VERSION AK010500.1 GI:12845991
 KEYWORDS CAP trapper.
 SOURCE Mus musculus (strain:C57BL/6J) ES cells cDNA to mRNA,
 clone.lib:RIKEN full-length enriched mouse cDNA library
 clone:2410015K23.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 845)
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Methods in enzymology. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2 (bases 1 to 845)

WashU-HMI Mouse EST Project
Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL: contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:462548
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 507.

FEATURES

Source

Location/Qualifiers
1. 612
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:761628"
/clone_lib="Barstead mouse pooled organs MPLRB4"
/sex="mixed"
/tissue_type="pooled organs"
/dev_stage="7 day"
/lab_host="DH10B"
/note="Organ: pooled; Vector: p77T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTACGAATCTGAAGTGGAGCGCGCCCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[CTTGATTCGCTACC]. digested with Not I and cloned into
the Not I and Eco RI sites of the modified p77T3 vector.
Library constructed by Bob Barstead."
168 a 167 c 175 g 102 t

BASE COUNT

ORIGIN

Query Match 84.3%; Score 590.4; DB 10; Length 612;
Best Local Similarity 99.0%; Pred. No. 1.9e-129;
Matches 594; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 14 cagcagctcgagctctctgcccgcgagcggcgtgacgagccgagcagcggac 73
|||||
Db 13 CACACGCTCGAGCTCTCTGCGCGCGCGCGGCTGCGAGCGCGCCGCGAGCGGGAC 72
|||||
QY 74 agattgactggaagcgcgagctccagcgcgagcgggaattgacagagagactacgcgc 133
|||||
Db 73 AGATTGACTGGAAGCGCGAGAGTCCAGCGCGCACGGGAATTGACAGGAGACTACCGCG 132
|||||
QY 134 aagggataggccagatagcaccgaggaacaaatctcatatggtgccaattgccacca 193
|||||
Db 133 AAGGGATAGGCCCGCAGAAATAGCAACCAAGCAAAATAATCTCATGTCGCAATGCCACCA 192
|||||
QY 194 ggaacacgaagctgagcagcgcctgcagaaatgacaggaacacgcgcctgtggagg 253
|||||
Db 193 GGAACACGAAGAGATGAGCAGCGCCCTGCAGAAATGACAGGAGAGCGCCCTGTGGAGG 252
|||||
QY 254 agtgaggggccacagcctgctgcaacacacacacacacacacacacacacacacacaa 313
|||||
Db 253 AGTGAGGGCCACCGCTGCTGCANACACACACACACACACACACACACACATACACCA 312
|||||
QY 314 ccaccaccgaagaggccagc 373
|||||
Db 313 CCACCACCGAAGAGCGCAGCTCGCGGACTTGCGCCCTAACTTCGATGGGCCCATTCCCA 372
|||||
QY 374 cagcagatgaatgacgggttggtgagatgagatgagatgatgatgaatgattgaggga 433
|||||
Db 373 CAGCGAGATGAATGACGGGTGGGTGGAGATGAGATGATGATGAATGTTTCATGGAGGA 432
|||||
QY 434 gatgagagatcccgagagaagcttagggagctcacagctgagaaaattgtctacgcactct 493
|||||
Db 433 GATGAGAGATCCGGGAGAAGCTTAGGGAGCTACAGCTGAGAAATGCTACGCACTCT 492
|||||
QY 494 tatgggggagctgtctaaaccacacacacacacacacacacacacacacacacacacac 553
|||||
Db 493 TATGGGGAGCTGTCTAAACCACACACACACACACATGATGAATCTGCGCTTATGCTTGACT 552
|||||

QY 554 tcggtcattcccccccgagatccatcactgtgactcccgctgtagccctttccctcgact 613
|||||
Db 553 TCGGTCATTCCCGCCCTGAGATCCATCTACTGCTGAGCTCCGCTGTAGCCCTTTCTCGCAT 612
|||||

RESULT 8
AK003294 LOCUS 826 bp mRNA HTC 05-JUL-2001
Mus musculus 18 days embryo cDNA, RIKEN full-length enriched
library, clone:1110002F04, full insert sequence.
AK003294 DEFINITION
AK003294 ACCESSION
AK003294 VERSION
AK003294 KEYWORDS
AK003294 SOURCE
Mus musculus (strain:C57BL/6J) 18 days embryo cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library
clone:1110002F04.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 826)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Methods in enzymology. 303, 19-44 (1999)
99279253
10349636
2 (bases 1 to 826)
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome research. 10 (10), 1617-1630 (2000)
20499374
11042159
3 (bases 1 to 826)
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujisaki, Y., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome research. 10 (11), 1757-1771 (2000)
20530913
11076861
4 (bases 1 to 826)
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 (bases 1 to 826)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M.,
Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiroka, T., Hori, F.,
Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J.,
Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T.,
Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y.,
Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H.,
Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Tanaka, T., Teijima, Y., Toya, T., Yamamura, T., Yasunishi, A.,
Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for

479 cgtcttaccgca-tccttatcgggggagctgtctaacccaccgacaccatgatgaa-ttc 536


```
|||||
Db 1 CGCGGGGGCTGGCAGCGGGCCCGAGCGAGCGGACAGATTGACTGGAAGCCGAGAGTC 60
QY 97 caggcgagcgggaattgacagagagactacgcgaaggagatagccagagaatagcaa 156
|||||
Db 61 CAGCGCGCACGGGGAAATTGACAGGAGGACTACCCCGCAAGGGATAGGCCCCAGAAATAGCAA 120
QY 157 ccaggaaacaaatctcatcatgccaatgtccaccaggaacgaagagctggagcagc 216
|||||
Db 121 CCAGGAACAANAATCTCATCATGCGCCAATGTCCACCAGGAACGAGAGATGGAGCAGC 180
QY 217 ccctgcagaatgacacaggaaacccgcctgtgtggagaggtgagggccaccagccctgctg 276
|||||
Db 181 CCCTGCAGAATGACAGGAAGACCGCCCTGTGGGAGGAGTGTAGGGGCCACCCAGCCTGCTG 240
QY 277 caaacaacaacaacaaccacaataaaccacaaccacaaccagaagagccaggttc 336
|||||
Db 241 CAAACAACAACAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
QY 337 gccgacctgccccctaaacttcctgatggccattcccaacaggcagatgaatgacgggttg 396
|||||
Db 301 GCCGACTTGCCCTTAACCTCCGATGGCCATTCCCAACAGGCAGATGAATGACGGGTGG 360
QY 397 gtggagatggagatgatggaatgttcctatggagagagatgagagagatccggagaaagc 456
|||||
Db 361 GTGGAGATGGAGATGATATGGAATGTTCATGGAGGAGATGAGAGAGATCCGGAGAAAGC 420
QY 457 ttaggagctacagctgagaaattgtctacgcatccttatggggagactgtctaacacc 516
|||||
Db 421 TTAGGGAGCTACAGCTGAGAAATTGTCTACGCATCCTTATGGGGGAGCTGTCTAACCAAC 480
QY 517 acgatcaccatgatgaattctgccttatgccttgacttcggtcattccccctgagatcc 576
|||||
Db 481 ACCATCACCATGATGAATCTGCGCTTATGCCCTTGACTTCGGTCATTGG -CCCTGAGATCC 539
QY 577 atactgtgactcc 589
|||||
Db 540 ATACTGTGACTCC 552
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 16:33:03 ; Search time 1319.57 Seconds
(without alignments)
11139.225 Million cell updates/sec

Title: US-09-327-750D-29

Perfect score: 891

Sequence: 1 accccatccccctctat.....aataagcaatttaaaagc 891

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.om.*

20: em.or.*

21: em.ov.*

22: em.pat.*

23: em.ph.*

24: em.pl.*

25: em.ro.*

26: em.sts.*

27: em.sy.*

28: em.un.*

29: em.vi.*

30: em.htgo_hum.*

31: em.htgo_inv.*

32: em.htgo_rod.*

33: em.htg_hum.*

34: em.htg_inv.*

35: em.htg_rod.*

36: em.htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Match	Length	DB	ID	Description
1	889.4	99.8	891	9	HUMOGC	M3188 Human unkno
2	889.4	99.8	891	9	AF187064	AF187064 Homo sapi
3	609.4	68.4	793	9	BC003190	BC003190 Homo sapi
4	363	40.7	854	10	AF097440	AF097440 Mus muscu
c 5	354.2	39.8	421	11	G72708	G72708 MARC 4953-4
6	312.2	35.0	700	10	AF187066	AF187066 Mus muscu
7	264.4	29.7	519	10	AF187065	AF187065 Rattus no
8	245.4	27.5	98274	9	HS71487	Z99755 Human DNA s
9	242.4	27.2	244	11	GL3550	GL3550 human STS S
10	185.2	20.8	45678	9	HSV351F8	Z70719 Human DNA s
11	158.6	17.8	69648	9	HS635G19	AL035494 Human DNA
12	141.2	15.8	1229	9	AK000959	AK000959 Homo sapi
13	141.2	15.8	1364	6	AX100231	AX100231 Sequence
14	141	15.8	43952	9	HSU10564	Z92846 Human DNA s
c 15	132.4	14.9	477	11	G35294	G35294 human STS S
c 16	122.2	13.7	35714	9	HS198P4	AL008708 Human DNA
c 17	117.2	13.2	154577	2	AL357046	AL357046 Homo sapi
c 18	117.2	13.2	154959	2	AC009282	AC009282 Homo sapi
c 19	117.2	13.2	185257	9	AL355305	AL355305 Homo DNA
20	110.8	12.4	31321	9	HSV870H8	Z70233 Human DNA s
21	110.8	12.4	41029	2	HS08081	AL022169 Homo sapi
22	109.8	12.3	642	9	AF183416	AF183416 Homo sapi
23	109.8	12.3	828	9	AF220189	AF220189 Homo sapi
c 24	109.2	12.3	40584	9	AL133348	AL133348 Human DNA
25	108.2	12.1	791	9	AF237783	AF237783 Homo sapi
26	104	11.7	898	6	AX078272	AX078272 Sequence
c 27	102.4	11.5	744	9	AF251053	AF251053 Homo sapi
c 28	77	8.6	43952	9	HSU10564	Z92846 Human DNA s
c 29	74.6	8.4	1163	9	BC005988	BC005988 Homo sapi
c 30	71.6	8.0	504	11	G24641	G24641 human STS W
c 31	66.4	7.5	785	10	AF097439	AF097439 Mus muscu
c 32	62.4	7.0	7218	6	I66494	I66494 Sequence 14
33	59.6	6.7	40822	9	HSCU177E8	Z68694 Human DNA s
34	59.4	6.7	138872	2	AC027560	AC027560 Homo sapi
c 35	59.4	6.7	176039	2	AC073065	AC073065 Homo sapi
c 36	58.6	6.6	1208	9	AK026349	AK026349 Homo sapi
c 37	58.6	6.6	91480	9	HSDJ823F3	AL079333 Human DNA
c 38	58.4	6.6	156275	2	AL513285	AL513285 Homo sapi
c 39	57	6.4	196367	9	AC020740	AC020740 Homo sapi
c 40	56.6	6.4	2269	10	AF097437	AF097437 Mus muscu
c 41	56.6	6.4	186896	2	AC037486	AC037486 Homo sapi
c 42	56.2	6.3	127606	2	AL356298	AL356298 Homo sapi
c 43	56.2	6.3	194019	2	AC053464	AC053464 Homo sapi
c 44	56.2	6.3	197105	2	AL355355	AL355355 Homo sapi
45	56.2	6.3	199891	9	CNS00M80	AL079303 Human chr

ALIGNMENTS

RESULT 1

HUMOGC

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

HUMOGC 891 bp mRNA 07-MAR-1995
Human unknown protein from clone PHGR74 mRNA, complete cds.

M38188 X58942

M38188.1 GI:189378

Human ovarian granulosa cell line, cDNA to mRNA.

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 891)

Rapp,G., Freudenstein,J., Klaudiny,J., Mucha,J., Wempe,F.,

Zimmer,M. and Scheit,K.H.

Characterization of three abundant mRNAs from human ovarian

granulosa cells

DNA Cell Biol. 9 (7), 479-485 (1990)

91025550

Draft entry and computer-readable sequence for [DNA 9, 479-485

(1990)] kindly submitted

```
by K.H.Scheit, 27-AUG-1990.
FEATURES             Location/Qualifiers
     source           1..891
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="pHGR74"
                     /cell_type="granulosa"
                     /tissue_type="ovary"
mRNA                 <1..891
                     /note="protein of unknown function"
CDS                  312..647
                     /codon_start=1
                     /note="protein of unknown function"
                     /protein_id="AA63232.1"
                     /db_xref="GI:189379"
                     /translation="MANIHQENEMEOPMQNGEEDRPLGGEGHQPAGNRRGQARRLA
PNFRWAIPNRQINDMGDDDDMEIFMEEMREIRKRLRELQRLNCLRLMGLSNNHHD
HHDEFCLMP"
CDS                  361..534
                     /note="protein of unknown function"
                     /codon_start=1
                     /protein_id="AA63233.1"
                     /db_xref="GI:189380"
                     /translation="MERKKTALWEEVKATSLQIBIDGDRDLADPLIFDGPYPIGRSMNGW
VEMEMIWKYSWRR"
BASE COUNT          251 a 182 c 224 g 234 t
ORIGIN

Query Match          99.8%; Score 889.4; DB 9; Length 891;
Best Local Similarity 99.9%; Pred. No. 3.8e-235;
Matches 890; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1  accccatccccactctactataccggtctctccatttggctgctgcaaaagctctggaaag 60
    |||||||
DB 1  ACCCCATCCCCACTCTCTATACCGGCTCCCATTTTGGTCCCTGCAAGCTCTGGGAAAG 60
    |||||||
QY 61  aatccccgggaacgaaataatggtgttggggaaggaggttaaggaggaagactga 120
    |||||||
DB 61  AATCCCGGGNACCAAAATGTTGGTGTGGGGAGGAGGTAAGGGAGAGAAAGCTGGA 120
    |||||||
QY 121 gggaggggttattggagggcccttagagagcgcggaacttctaagggtgggaaaaa 180
    |||||||
DB 121 GCGAGGGGCTTTAATTTGGAGGCCCTTAGAGAGCGCGGAACTTCTAAGGTGGGAAAAA 180
    |||||||
QY 181 acgaataataaatctttgatatacagggtctgtaactcctgctgctgagaccacaagc 240
    |||||||
DB 181 ACGAAATTAATAAATCCTTTGATATCAGGGCTCTGAATCCTGCTGTCAGAGCACCACAGC 240
    |||||||
QY 241 attcagtcctctctctgcttctgttcttacttctgttctaaagaaaaaacacagaaaaa 300
    |||||||
DB 241 ATTCACTCTCTCTCTTCTGCTTTGCTTTACTTGTGTTCAAAGAAAAAACACAGAAAAA 300
    |||||||
QY 301 aaatctcatctatggcaaatattcaccaggaacgaagagatggagcagcctatgcaga 360
    |||||||
DB 301 AAAATCTCATGCGCAATATTTCACGAGAAACGAGAGATGGAGAGCCTATGCCAGA 360
    |||||||
QY 361 atggagagaagaccgcttgggaggggtgaaggccaccagcctgcaggaaatcgac 420
    |||||||
DB 361 ATGGAGAGAAAGACCGCCCTTTGGGAGGAGGTGAAGCCACCAGCCTGCAGGAAATCGAC 420
    |||||||
QY 421 ggggagagctcgccgactgcccctaatcttcgattggcccatcccaataggcagatca 480
    |||||||
DB 421 GGGACAGAGCTCCCGACTTGCCCCCTTAATTTTCGATGGGCCATACCAATATGGCAGATCA 480
    |||||||
QY 481 atgatgggatgggtgagatgagatgatggaataatttcattgaggagagatgagagaa 540
    |||||||
DB 481 ATGATGGGATGGGTGAGATGGAGATGATGGAATATTTCATGGAGGAGATGAGAGAA 540
    |||||||
QY 541 tcagaagaaacttagggagctgcagttgagggaattgtcgtatccttatggggagc 600
    |||||||
DB 541 TCAGAAGAAACTTAGGGAGCTGCAGTTGAGGAATTTGCTCGGTATCTCTTATGCGGGAGC 600
    |||||||
```



```
QY 521 catggaggagatgagagaaatcagaagaaacttagggagctgcagttgaggaattgtct 580
|||||
Db 405 CATGGAGGAGATGAGAGAAATCAGAAGAAACTTAGGGAGCTGCAGTTGAGGAATTGTCT 464

QY 581 gcgtatccattatggggagctcttaatacacatgaccatcatgatgaattttgacctat 640
|||||
Db 465 GCGTATCCCTATGATGGGGAGCTCTTAATCACCATGACCATCATGATGAATTTTGCCTTAT 524

QY 641 gccttgactcctgccatttatcatgatgagattaactactgtgattcccgctgtttcttttc 700
|||||
Db 525 GCCTTGACTCCCGCCATTTATCATGAGATTAATACATGTGATGCCCTGTTCTTTTTC 584

QY 701 ctgtcattttcctaataatgoccttactatccgtttgctgtgtaacocctatgtatttcca 760
|||||
Db 585 CTGTGCAATTTTCCCTAATATGCCTTTACTGATCCGTTTGTGCTGAACCCATGATTTATTTCCA 644

QY 761 tgtgtcaagtgggtctgtgtgctgagcttctatttgaagattgcctttgacctcagttc 820
|||||
Db 645 TGTGTCAAGTGGGTCTGTGTGTTGCCAGCTTCTATTTTGAAGATTGCCCTTTCACCTCAGTGT 704

QY 821 aaqtttctgtcagcagtagtttcacccatttgcattgagaaattttaaagccaaataagca 880
|||||
Db 705 AAGTTTCTGTACGACGACTAGTTTACCCATTTCGATGGAATAATTAAGCTAATAAAGCA 764

QY 881 atttaaaagc 891
|||||
Db 765 ATTTAAAGC 775

RESULT 4
AF097440 854 bp mRNA ROD 13-APR-1999
LOCUS Mus musculus brain expressed X-linked protein 3 (Bex3) mRNA,
DEFINITION complete cds.
ACCESSION AF097440
VERSION AF097440.1 GI:4580593
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 854)
Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
Hum Mol Genet. 8 (4), 611-619 (1999)
JOURNAL Hum Mol Genet. 8 (4), 611-619 (1999)
MEDLINE 99172070
REMARK Erratum: [[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
2 (bases 1 to 854)
Brown,A.L. and Kay,G.F.
Direct Submission
Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
FEATURES
Location/Qualifiers
1..854
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/db_xref="dbEST:AA272375"
/chromosome="X"
/map="near p1p"
/tissue_type="pooled organs"
1..854
/gene="Bex3"
172..546
/gene="Bex3"
/codon_start=1
/product="brain expressed X-linked protein 3"
/protein_id="AAD24431.1"
/db_xref="GI:4580594"
1..854
/translation="MANVHDSNEEMEQPLONGQEDRPVGGEGHQPAAANNHNNHNN
HNHRRGQARLAPNFRWAIPIRQNDGLGGDDMMFMEMREIRRLKRLQLRNC
```

```
LRILMAGELSNHHDDHDEFCLMP"
BASE COUNT 237 a 212 c 228 g 177 t
ORIGIN
Query Match 40.7%; Score 363; DB 10; Length 854;
Best Local Similarity 78.8%; Pred. No. 1.6e-89;
Matches 516; Conservative 0; Mismatches 90; Indels 49; Gaps 5;
QY 284 aaaaacacagaaaaaataatctcatcatggcaataatctcaccaggaaaaacgaagagat 343
|||||
Db 144 AATAGCAACACGAGAAACAAATCTCATATGCGCAATGTCCACAGGAAACGAAGAT 203

QY 344 gtagcagcctatgcagaatgcagagagagaccccttctggagagagtgagagccacca 403
|||||
Db 204 GGAGAGCCCTTGCAGATGGACAGAACCCCTGTGGGAGGAGGTGAGGGGCCACCA 263

QY 404 gcctgcagaa-----atcgacggg 424
|||||
Db 264 GCCTGCTGCAACAAACAACAACAACCAACCAATACCAACCAACCCGGAAGAGG 323

QY 425 acaggctgcgcagacttgccttaattttcgatgggccaatacccaataggcagatcaatga 484
|||||
Db 324 CCAGGCTCGCGGACTTGCCCTTAACCTTCGGATGGGCCATTTCCCAACAGCAGATGAATGA 383

QY 485 tgggatgggtgagatggagatgatatgaaatattcatggagagagatgagaaatcag 544
|||||
Db 384 CGGGTGGGTGGAGATGGAGATGATGAAATGTTTCATGGAGGAGATGAGAGATCCG 443

QY 545 aagaaaacttagggagctgcagttgaggaattgtctgcgtatccttatggggagctctc 604
|||||
Db 444 GAGAAAGCTTAGGGAGCTACAGCTCAGAAATTTGCTACGCAATCCTTATGGGGAGCTGTC 503

QY 605 taatcacatgaccatcatgatgaattttgccttgccttgccttgccttgccttgccttgc 663
|||||
Db 504 TAACCACCAACGATCACCATGATGAATTTCTGCTTATGCTTTCGCTTTCGCTTTCGCTTTC 563

QY 664 tgagattaactgtgattcccgctgtttcttcttcttcttcttcttcttcttcttcttctt 723
|||||
Db 564 TGAGATCCATATCATGTGACTCCGCTGTGACCCCTTTTCTCGCATTTTCTCGCATTTTCTCG 623

QY 724 tactgatcgtttgtgtgaacctatgtattt-----ccatgtgtcaagtgggtcttg 778
|||||
Db 624 TAATGACCCGTTTGGTGGAGCCTTGTGTATTATTCATGCCATGTGCCAGGTGGGGCTTG 683

QY 779 tgttgcagctctctatttgaagattgccttgccttgccttgccttgccttgccttgccttgc 837
|||||
Db 684 TGTGCCAGCTTCTAATTTGGAGATTGCCCTTCCCACTCAGTGTGAGTTTCTGTCAACAGTG 743

QY 838 -agtttcacccatttgcagtaaaa--aatttaagccaaataaagcaatttaaaaa 889
|||||
Db 744 CAGTGTGCCCATTAGCATGGAAGCAATGTCAAGTCAATAAAGCAACTGGGAATA 798

RESULT 5
G72708/c 5
LOCUS G72708 421 bp DNA STS 08-AUG-2001
DEFINITION MARC 4953-4954:991939031:1 SCF - porcine spleen Sus scrofa STS
genomic, sequence tagged site.
ACCESSION G72708
VERSION G72708.1 GI:15146738
KEYWORDS STS.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 421)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L. and
Keele,J.W.
REFERENCE Single nucleotide polymorphism (SNP) discovery in expressed porcine
genes
JOURNAL Unpublished (2001)
COMMENT
```

Contact: Preking BA
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4278
Fax: 402 762 4173
Email: freking@email.marc.usda.gov
Primer A: GCAAAATGGTGAAACTCTACT
Primer B: CCAGCAGCAAAATAGACG
STS size: 500
PCR profile:

Hotstart: 95 degrees for 15 minutes
Denature: 95 degrees for 30 seconds
Anneal: 56 degrees
Extension: 68 degrees for 2 minutes
Cycles: 32 to 45

Protocol:
Template: 50-200 ng genomic DNA
Primer: each 20 pmol
dNTPs: each 88 uM
Taq-Polymerase: 0.25 units (Qiagen HotStar)

Buffer: Commercially supplied Qiagen HotStar buffer

The STS is derived from PCR amplicons generated from genomic DNA, sequenced from each end using the amplification primers. The sequence does not necessarily represent the entire amplicon. Sequence derived from PolyPhred was trimmed from each end of each unique contig until five consecutive bases exceeded a quality score threshold of 20, and the next 10 bases averaged a quality score of 20 or greater. Amplicon size was estimated by agarose gel electrophoresis.

FEATURES

Source

Location/Qualifiers
1..421
/organism="Sus scrofa"
/strain="white composite, duroc, meishan, minzhu, fengjing, crossbreds"
/db_xref="taxon:9823"
/sex="male and female"
/clone_lib="SCF - porcine spleen"
/dev_stage="adult"
/note="Organ: spleen"
<1..>421

STS

BASE COUNT 130 a 101 c 86 g 102 t 2 others
ORIGIN

Query Match 39.8%; Score 354.2; DB 11; Length 421;
Best Local Similarity 92.1%; Pred. No. 3.8e-87;
Matches 384; Conservative 1; Mismatches 29; Indels 3; Gaps 1;
QY 414 aatcgacggggacaggctcgccgactgcccctaatttcgatggccatacccaatagg 473
|||||
Db 421 AATAGACGGGGACAAGCTCGCGACTGCGCCCTAAATTCGATGGGCCATACCAATAG 362
474 cagatcaatgatggatgggtggagatggagatgatgaaatattcattgatggagagatg 533
|||||
Db 361 CAGATCAATGATGGATGGTGGAGATGGAGATGATGGAATGTTTCATGGAGGAGATG 302
534 agagaaatcagagaaataattagggagctgcagttgaggaattgtctgcgtatccctatg 593
|||||
Db 301 AGAGAAATCAGGAGAAACCTTAGGAGCTGCAGTTGAGGAATGTCTGCGTATCCTTATG 242
594 gggagactctctaataccacatgacacatgatgaattttgccttatgcttgactccctg 653
|||||
Db 241 GGGAGACTCTCTAATCACCATGACCATCATGATGAATTTTGGCTTATGCTTGACTTCCTG 182
654 ccatttatcatgagataactgtgattccgcgtgtt---ttcttttcccttgcaatttt 710
|||||
Db 181 CCATTTTCATGAGGTTAATACTGTGATYCCCACTGTTGYCTCTTTTTCCTTACATTTT 122
711 cctaataatgcctttactgatccggtttgctgtgaacccctatggttatttccatgtgtcaagt 770
|||||

Db 121 CCTGATATGCCTTTACTGATCCATTGCTCTGAACCTTATGTAAATTTCCATGTGTGAGT 62
QY 771 gggcttctgtgtccagcttcttatttgaagattgccttgcctcagtgtaagtcc 827
|||||
Db 61 GGGTCTCTGTGTACCAGCTTCTAATTTGGAGATTGCCTTGGCACCCCAAAAGTTCTTCTGTC 5
RESULT 6
LOCUS AF187066 700 bp mRNA ROD 11-JUN-2000
DEFINITION Mus musculus p75NTR-associated cell death executor (Nade) mRNA, complete cds.
ACCESSION AF187066
VERSION AF187066.1 GI:8452897
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 700)
AUTHORS Mukai,J., Hachiya,T., Shoji-Hoshino,S., Kimura,M.T., Nadano,D., Suvanto,P., Hanaoka,T., Li,Y., Irie,S., Greene,L.A. and Sato,T.A.
TITLE Nade, a p75NTR-associated cell death executor, is involved in signal transduction mediated by the common neurotrophin receptor p75NTR
J. Biol. Chem. 275 (23), 17566-17570 (2000)
MEDLINE 20298829
REFERENCE 2 (bases 1 to 700)
AUTHORS Mukai,J., Hachiya,T., Hoshino,S., Kimura,M., Nadano,D., Suvanto,P., Hanaoka,T., Li,Y., Irie,S. and Sato,T.
TITLE Direct Submission
JOURNAL Submitted (17-SEP-1999) Otolaryngology/Pathology, Columbia University, 630 West 168th St., P&S 11-451, New York, NY 10032, USA
FEATURES
Location/Qualifiers
1..700
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
1..700
/gene="Nade"
/gene="Nade"
177..551
/note="Involved in the common neurotrophin receptor p75NTR-mediated signal transduction; Nade"
/codon_start=1
/product="p75NTR-associated cell death executor"
/protein_id="AAF75131.1"
/db_xref="GI:8452898"
/translation="MANYHQENEEMEQPLQNGQEDRPVGGEGHQPAAANNHNNHNNHNNHRRGGARLAPNFRWAI PNRMNDGLGGDDMEMEEMREIRKRLRELQLRNC
LRLMGLSNHHDHDEFCLMP"
BASE COUNT 178 a 187 c 203 g 132 t
ORIGIN
Query Match 35.0%; Score 312.2; DB 10; Length 700;
Best Local Similarity 78.5%; Pred. No. 1.7e-75;
Matches 431; Conservative 0; Mismatches 73; Indels 45; Gaps 3;
QY 284 aaaaacacagaaaaaaatctcatcgtgcaaatatttcaccaggaagaaagagat 343
|||||
Db 149 AATAGCAACCGAGAAACAAATCTCATGTGCCAATGTCCACCAGAAACGACAGAT 208
344 ggaagcgcctatgcagaaatggagaggaagacgcgcctttggaggagggtgaaggccacca 403
|||||
Db 209 GGAGCAGCCCTTCAGAAATGGACGAGAACGCCCTGTGGGAGGAGTGTGAGGGCCACCA 268
404 gctgcagaa-----atcgacgggg 424
|||||
Db 269 GCTGTGCAACAACAACAACAACCAACCAACCAACCAACCAACCAACCAACCAAGAGG 328
425 acaggctcgcgacttgcctcctaatttctgatggcgccatacccaataggcagatcaatga 484
|||||


```

RESULT 10
HSV351F8      45678 bp      DNA      PRI      23-NOV-1999
LOCUS
DEFINITION    Human DNA sequence from cosmid V351F8, between markers DXS366 and
               DXS87 on chromosome X contains ESTs.
ACCESSION     Z70719
VERSION       Z70719.1 GI:1261915
KEYWORDS      X.
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 45678)
AUTHORS       Whiteley, M.
TITLE         Direct Submission
JOURNAL       Submitted (09-Apr-1995) Sanger Centre, Hinxton, Cambridgeshire,
               CB10 1RQ, UK. E-mail enquiries: humquery@sanger.ac.uk
COMMENT       IMPORTANT: This sequence is the entire insert of clone V351F8. The
               true left end of clone V351F8 is at 1 in this sequence. The true
               right end of clone V351F8 is at 45678.
               V351F8 is from the human chromosome X-specific cosmid library.
FEATURES      Location/Qualifiers
               1..45678
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="X"
               /map="X"
               /clone="GHC-351F8"
               /clone_lib="SCCV"
               1..70
repeat_region
               /partial
               /note="Alu repeat: matches 80. .1 of consensus"
               642..930
repeat_region
               /note="Alu repeat: matches 1. .308 of consensus"
               1406..5179
repeat_region
               /note="L1 element fragment"
               5243..5276
repeat_region
               /note="17 copies of 2 mer 82 % conserved"
               5279..5567
               /partial
               /note="Alu repeat: matches 304. .1 of consensus"
               9975..10369
repeat_region
               /note="MSTA element fragment"
               10399..12017
repeat_region
               /note="MSTAR element fragment"
               11318..11978
repeat_region
               /note="THE1BR element fragment"
               11715..11846
repeat_region
               /note="MLTIR element fragment"
               12074..12170
repeat_region
               /note="MSTA element fragment"
               12227..12486
repeat_region
               /note="MSTA element fragment"
               12437..12502
repeat_region
               /note="MSTC element fragment"
               12756..13059
               /partial
               /note="Alu repeat: matches 308. .1 of consensus"
               14478..14650
repeat_region
               /note="THE1B element fragment"
               14491..14536
repeat_region
               /note="MSTA element fragment"
               14715..14831
repeat_region
               /note="THE1B element fragment"
               15099..15533
repeat_region
               /note="L1 element fragment"
               15534..15626
repeat_region
               /note="31 copies of 3 mer 85 % conserved"
               15566..15619
repeat_region
               /note="3 copies of 18 mer 98 % conserved"
               16352..16436
repeat_region
               /note="L1 element fragment"
               16728..16763

```

```

repeat_region  /note="9 copies of 4 mer 94 % conserved"
               16764..17054
               /partial
               /note="Alu repeat: matches 308. .1 of consensus"
               19420..19449
repeat_region  /note="15 copies of 2 mer 87 % conserved"
               complement(19537..20013)
misc_feature   /note="match: 3' EST N51315 clone 283089"
               20071..20114
repeat_region  /note="22 copies of 2 mer 98 % conserved"
               20073..20112
repeat_region  /note="10 copies of 4 mer 100 % conserved"
               20073..20126
repeat_region  /note="3 copies of 18 mer 87 % conserved"
               23088..23300
repeat_region  /note="L1 element fragment"
               23311..23490
repeat_region  /note="MLTIA element fragment"
               23318..23489
repeat_region  /note="MLTIB element fragment"
               23362..23491
repeat_region  /note="MLTIC element fragment"
               23507..24407
repeat_region  /note="L1 element fragment"
               24376..24462
repeat_region  /note="MSTC element fragment"
               24379..24491
repeat_region  /note="MSTA element fragment"
               24688..24738
repeat_region  /note="MSTA element fragment"
               24810..25016
repeat_region  /note="L1 element fragment"
               25254..25421
repeat_region  /note="L1 element fragment"
               25439..25696
repeat_region  /note="MER25 element fragment"
               26284..26343
repeat_region  /note="MLTID element fragment"
               26428..26491
repeat_region  /note="MLTID element fragment"
               26465..26518
repeat_region  /note="MLTIE element fragment"
               27446..27667
repeat_region  /note="L1 element fragment"
               28954..29247
               /partial
               /note="Alu repeat: matches 308. .1 of consensus"
               30160..30325
repeat_region  /note="2 copies of 83 mer 98 % conserved"
               32588..32875
misc_feature   /note="match: 5' EST H68599 clone 239077"
               complement(32825..33230)
misc_feature   /note="match: 3' EST H68239 clone 289077"
               35551..35802
               /partial
               /note="Alu repeat: matches 1. .260 of consensus"
               35804..35839
repeat_region  /note="18 copies of 2 mer 83 % conserved"
               38625..38705
repeat_region  /note="MLTIR element fragment"
               38747..38839
repeat_region  /note="MLTIR element fragment"
               38949..39032
repeat_region  /note="MLTIE element fragment"
               38965..39076
repeat_region  /note="MLTID element fragment"
               9135 c 9327 g 12956 t
BASE COUNT    14260 a
ORIGIN

```

Query Match 20.8%; Score 185.2; DB 9;
 Best Local Similarity 64.4%; Pred. No. 5.2e-40;
 Matches 345; Conservative 0; Mismatches 178; Indels 13; Gaps 4;


```

RESULT 12.
AK000959
LOCUS Homo sapiens cDNA FLJ10097 fis, clone HEMBA1002458, weakly similar
        to OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74.
DEFINITION
ACCESSION AK000959
VERSION 1
KEYWORDS ologo capping; fis (full insert sequence).
SOURCE Homo sapiens embryo, 10 weeks whole embryo, mainly head cDNA to
        mRNA, clone_lib:HEMBA1 clone:HEMBA1002458.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K.,
Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y. and Oshima,A.
WAKAMATSU A., NAKAMURA Y., NAGAHARI K., MASUHO Y. and OSHIMA A.
NEDO human cDNA sequencing project
Unpublished (2000)
2 (bases 1 to 1229)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (16-FEB-2000) to the DDBJ/EMBL/GenBank databases. Takao
Isogai, Helix Research Institute, Genomics Laboratory: 1532-3 Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp,
Tel:81-438-52-3951, Fax:81-438-52-3952)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5', 3' end one pass sequencing and clone selection;
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
Location/Qualifiers
1..1229
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEMBA1002458"
/clone_lib="HEMBA1"
/dev_stage="embryo, 10 weeks"
/tissue_type="whole embryo, mainly head"
/Note="Cloning vector: pME18SFL3"
191..553
/Note="unnamed protein product"
/codon_start=1
/protein_id="BAA91443.1"
/db_xref="GI:7021946"
/translation="MESKEELAAANNLNGENAQQNEGEQAPONEEESRHLGGEGQ
KPGNIRGRVRRLLVPNFRWAIIPNRHIEHNEARDVVERFVGOMMEIKRKTREQQMRHY
MRFTQPEPDNRHDFCLIP"
BASE COUNT 351 a 243 c 296 g 339 t
ORIGIN

Query Match 15.8%; Score 141.2; DB 9; Length 1229;
Best Local Similarity 68.4%; Pred. NO. 3.9e-28;
Matches 264; Conservative 0; Mismatches 98; Indels 24; Gaps 4;

QY 283 aaaaaccaccagaaaaaaatctcatcgatggcaaatattcaccaggagaaacgaagaga 342
Db 201 AAGAGAACTAGCGCAACAATCTCAACGGGGAAATGCCCAACAAGAAACGAAGGAG 260
QY 343 ttgagcag---cctatgcagatggaggaagaccgcccctttggagaggtgaagcc 399
Db 261 GGGACAGCCCCCAGCGAGATGAAGAAGATCCGCCATTTGGAGGGGTGAAGGCC 320
QY 400 accagcctgcaggaat---cgacggggacaggtcgccgaccttgcacctatttcgat 456
Db 321 AGAAGCCTGGAGGAATATCAGCGGGGGCGAGTTAGGCGACTTTCCTTAATTTTCGAT 380

Query Match 15.8%; Score 141.2; DB 6; Length 1364;
Best Local Similarity 68.4%; Pred. NO. 4e-28;
Matches 264; Conservative 0; Mismatches 98; Indels 24; Gaps 4;

QY 283 aaaaaccaccagaaaaaaatctcatcgatggcaaatattcaccaggagaaacgaagaga 342
Db 260 AAGAGAACTAGCGCAACAATCTCAACGGGGAAATGCCCAACAAGAAACGAAGGAG 319
QY 343 ttgagcag---cctatgcagatggaggaagaccgcccctttggagaggtgaagcc 399
Db 320 GGGACAGCCCCCAGCGAGATGAAGAAGATCCGCCATTTGGAGGGGTGAAGGCC 379
QY 400 accagcctgcaggaat---cgacggggacaggtcgccgaccttgcacctatttcgat 456
Db 380 AGAAGCCTGGAGGAATATCAGCGGGGGCGAGTTAGGCGACTTTCCTTAATTTTCGAT 439
QY 457 gggccataccataggcagatcaatgatggatggggtgagatggatgatgatgga 516
Db 440 GGGCCATACCTTAATAGGCATAT-----TGACCAATGAAGCGGAGATGATGAGAAA 493
QY 517 tttcatggagagatgagagaatcagaagaaacttagggagctgcagttgaggaatt 576
Db 494 GCTTTGTAGGCGAGATGATGGAATCAAGAGAAAGACTAGGCAACAGCAGATGAGGCACT 553
QY 577 gctcgctatcctctatgggggagctctctaaaccatgaccatcatgatgaatttgc 636
Db 554 ATATCGCTTCCAAA-----CTCCTGAACCTGACAACCATTTATGACTTTTGCC 601

```

```

QY 457 gggccataccataggcagatcaatgatggggtgagatggatgatgatgga 516
Db 381 GGGCCATACCTTAATAGGCATAT-----TGACCAATGAAGCGGAGATGATGATGAGAAA 434
QY 517 tttcatggagagatgagagaatcagaagaaacttagggagctgcagttgaggaatt 576
Db 435 GGTTTGTAGGCGAGATGATGGAATCAAGAGAAAGACTAGGCAACAGCAGATGAGGCACT 494
QY 577 gctcgctatcctctatgggggagctctctaaaccatgaccatcatgatgaatttgc 636
Db 495 ATATCGCTTCCAAA-----CTCCTGAACCTGACAACCATTTATGACTTTTGCC 542
QY 637 ttatgccttgactccttgccttatttc 662
Db 543 TCATACCTTGAATCCTTAAAGTTTTC 568

RESULT 13
AX100231
LOCUS AX100231 1364 bp DNA PAT 02-APR-2001
DEFINITION Sequence 32 from Patent WO0119860.
ACCESSION AX100231
VERSION AX100231.1 GI:13539115
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1364)
AUTHORS Tang,Y.T., Hillman,J.L., Yue,H., Reddy,R., Lal,P., Shah,P.,
Azlmzai,Y., Baughn,M.R., Lu,D.A., Bandman,O., Shih,L.L. and
Patterson,C.
TITLE Proteins associated with cell differentiation
JOURNAL Patent: WO 0119860-A 32 22-MAR-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..1364
/organism="Homo sapiens"
/db_xref="taxon:9606"
/Note="Incyte ID No: 1990956CB1"
BASE COUNT 411 a 269 c 322 g 362 t
ORIGIN

Query Match 15.8%; Score 141.2; DB 6; Length 1364;
Best Local Similarity 68.4%; Pred. NO. 4e-28;
Matches 264; Conservative 0; Mismatches 98; Indels 24; Gaps 4;

QY 283 aaaaaccaccagaaaaaaatctcatcgatggcaaatattcaccaggagaaacgaagaga 342
Db 260 AAGAGAACTAGCGCAACAATCTCAACGGGGAAATGCCCAACAAGAAACGAAGGAG 319
QY 343 ttgagcag---cctatgcagatggaggaagaccgcccctttggagaggtgaagcc 399
Db 320 GGGACAGCCCCCAGCGAGATGAAGAAGATCCGCCATTTGGAGGGGTGAAGGCC 379
QY 400 accagcctgcaggaat---cgacggggacaggtcgccgaccttgcacctatttcgat 456
Db 380 AGAAGCCTGGAGGAATATCAGCGGGGGCGAGTTAGGCGACTTTCCTTAATTTTCGAT 439
QY 457 gggccataccataggcagatcaatgatggatggggtgagatggatgatgatgga 516
Db 440 GGGCCATACCTTAATAGGCATAT-----TGACCAATGAAGCGGAGATGATGAGAAA 493
QY 517 tttcatggagagatgagagaatcagaagaaacttagggagctgcagttgaggaatt 576
Db 494 GCTTTGTAGGCGAGATGATGGAATCAAGAGAAAGACTAGGCAACAGCAGATGAGGCACT 553
QY 577 gctcgctatcctctatgggggagctctctaaaccatgaccatcatgatgaatttgc 636
Db 554 ATATCGCTTCCAAA-----CTCCTGAACCTGACAACCATTTATGACTTTTGCC 601

```


THIS PAGE BLANK (USPTO)

OM of: US-09-327-750D-30 to: GenEmbl: * out_format : pfs
Date: Mar 11, 2002 3:33 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=Cgnt2.1/USPTO.spool/US09327750/runat_l1032002_l01154_20319/app_query.fasta_1.1472
-DB=GenEmbl -QFW=fastap -SUFFIX=rge -GAP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -OGAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FCGAPOP=6.000
-DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFWT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=2000000000 -USER=US09327750_@CGNT_1_18673
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-30
Query length: 128
Database: GenEmbl: *
Database sequences: 1472140
Database length: -34134837
Search time (sec): 4557.230000

score_list:

Sequence	Strd	Orig	zScore	Escore	Len	Documentation
gb_ro:AF097438	+	688.00	1035.50	2.1e-49	835	! AF097438 Mus musculus brain exp
gb_ro:AF051347	+	688.00	1035.47	2.1e-49	838	! AF051347 Mus musculus REX-3 mRN
gb_ro:AF097437	+	683.00	1020.55	1.4e-48	2269	! AF097437 Mus musculus Bexl prc
gb_ro:BC003254	+	619.00	922.39	4.2e-43	2901	! BC003254 Mus musculus, Similar
gb_ro:AF097439	+	589.50	887.68	3.6e-41	785	! AF097439 Mus musculus brain exp
gb_pr:AF183416	+	453.50	684.45	7.6e-30	642	! AF183416 Homo sapiens ovarian c
gb_pr:AF220189	+	453.50	682.56	9.6e-30	828	! AF220189 Homo sapiens uncharact
gb_pr:HS1984	+	453.50	654.61	3.5e-28	35714	! AL008708 Human DNA sequence
gb_pr:AF237783	+	450.50	678.38	1.6e-29	791	! AF237783 Homo sapiens brain-exp
gb_pr:AF251053	+	450.00	678.08	1.7e-29	744	! AF251053 Homo sapiens X-linked
gb_pr:AX078272	+	450.00	676.69	2.0e-29	898	! AX078272 Sequence 76 from Paten
gb_pr:HSV870H8	+	450.00	650.32	6.0e-28	31321	! 270233 Human DNA sequence fr
gb_pr:AL133348	+	450.00	648.39	7.7e-28	40584	! AL133348 Human DNA sequence
gb_hgt:HSU80B1	+	450.00	648.31	7.8e-28	41029	! AL022169 Homo sapiens chromos
gb_sts:G24641	+	299.00	453.66	5.4e-17	504	! G24641 human STS WI-11354, sequ
gb_pr:HSU351F8	+	172.50	229.78	0.0002	45678	! 270719 Human DNA sequence fr
gb_pr:HMUGC	+	171.50	257.50	4.6e-06	891	! M38188 Human unknown protein fr
gb_pr:AF187064	+	170.50	256.86	5.0e-06	793	! AF187064 Homo sapiens p75NTR-a
gb_pr:BC003190	+	170.00	252.86	8.3e-06	1229	! BC003190 Homo sapiens, p75NTR-a
gb_pr:AK000959	+	170.00	252.86	8.3e-06	1229	! AK000959 Homo sapiens cdna FLJ
gb_pr:AX100231	+	170.00	252.08	9.2e-06	1364	! AX100231 Sequence 32 from Paten
gb_pr:HS635G19	+	170.00	222.88	0.0004	69648	! AL035494 Human DNA sequence
gb_ro:AF187066	+	161.00	243.49	2.8e-05	700	! AF187066 Mus musculus p75NTR-as
gb_ro:AF097440	+	161.00	242.01	3.3e-05	854	! AF097440 Mus musculus brain exp
gb_sts:G35294	+	159.00	240.32	2.8e-05	477	! G35294 human STS SHGC-37409, se
gb_ro:AF187065	+	135.00	206.57	0.0031	519	! AF187065 Rattus norvegicus p75N
gb_sts:G27208	+	130.50	201.35	0.0061	421	! G27208 MARC 4953-4954:991939031
gb_sts:G23964	+	100.50	157.11	1.79	372	! G23964 human STS WI-15922, sequ
gb_hgt:AC022779	+	96.50	104.49	1.5e+03	197886	! AC022779 Mus musculus clone
gb_hgt:AC023542	+	95.50	110.06	7.45e-84	76295	! AC023542 Homo sapiens chromos
gb_pl:AF271894	+	95.00	133.72	35.85	2845	! AF271894 zea mays lipoxigenase
gb_pr:AB040371	+	94.50	129.72	59.91	4408	! AB040371 Homo sapiens mRNA for
gb_hgt:AF001332	+	94.00	103.47	1.7e+03	136678	! AF001332 Homo sapiens chromo
gb_pr:AC079936	+	94.00	103.13	1.8e+03	143163	! AC079936 Genomic Sequence FC
gb_hgt:AP003440	+	94.00	102.21	2.0e+03	162095	! AP003440 Homo sapiens chromo
gb_hgt:AC087505	+	94.00	102.06	2.1e+03	165434	! AC087505 Homo sapiens chromo
gb_hgt:AP002893	+	94.00	100.62	2.5e+03	200841	! AP002893 Homo sapiens chromo
gb_hgt:AC028293	+	94.00	100.50	2.5e+03	204129	! AC028293 Homo sapiens chromo
gb_pr:A25464	+	93.50	140.49	15.06	844	! A25464 B.pertussis STR pilin ge
gb_ba:BPPIMX	+	93.50	140.48	15.08	845	! Y00556 Bordetella pertussis fin

gb_ba:BBFIMX + 93.50 137.19 22.97 1315 ! X74118 B.bronchiseptica fim
gb_pr:AF13843 - 93.50 137.19 22.97 1315 ! A73843 Sequence 1 from Pate
gb_pr:AF152500 - 93.00 132.03 44.54 2382 ! AF152500 Homo sapiens proto
gb_pr:AK024641 - 93.00 130.73 52.61 2837 ! AK024641 Homo sapiens cDNA:
gb_pr:AF217750 - 93.00 129.08 65.08 3547 ! AF217750 Homo sapiens proto

seq_name: gb_ro:AF097438

seq_documentation_block:
LOCUS AF097438 835 bp mRNA ROD 11-APR-1999
DEFINITION Mus musculus brain expressed X-linked protein 1 (Bexl) mRNA,
complete cds.

ACCESSION AF097438

VERSION AF097438.1 GI:4580589

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE 1 (bases 1 to 835)

JOURNAL Brown, A.L. and Kay, G.F.

MEDLINE Bexl, a gene with increased expression in parthenogenetic embryos,

REMARK is a member of a novel gene family on the mouse X chromosome

Hum. Mol. Genet. 8 (4), 611-619 (1999)

Erratum: [[published erratum appears in Hum Mol Genet 1999

May; 8(5):943]]

REFERENCE 2 (bases 1 to 835)

AUTHORS Brown, A.L. and Kay, G.F.

TITLE Direct Submission

JOURNAL Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of

Medical Research, Herston Rd, Brisbane, Qld 4029, Australia

FEATURES Location/Qualifiers

1..835

/organism="Mus musculus"

/db_xref="taxon:10090"

/chromosome="X"

/map="near plp"

/tissue-type="blastocyst"

1..835

/gene="Bexl"

205..591

/gene="Bexl"

/codon_start=1

/product="brain expressed X-linked protein 1"

/protein_id="AAD24429.1"

/db_xref="GI:4580590"

/translation="MESKDGQVKNLNMENDHOKKEEKPDQTIIRREPAVALISEAG

KNCAPRGRRFRVRQPIAHYRWDLMRVGEPOGRMREENVORFGDVRQLMEKLRR

QLSHSLRAVSTDPPHDHHDFCLMP"

BASE COUNT 238 a 184 c 246 g 167 t

ORIGIN

alignment_scores:

Quality: 688.00 Length: 128

Ratio: 5.375 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-327-750D-30 x AF097438 ..

Align seg 1/1 to: AF097438 from: 1 to: 835

1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspH1 17

205 ATGGAGTCCAAAGATCAAGCGTGAAATCTCAACATGGAGATGACCA 254

17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleArgArg 34

255 TCAGAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 304

34 luproAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50

```
|||||
305 AGCCAGCTGTGGCCTCATCTCCAGGCTGCAAAACTGTCGCCCTAGA 354
51 GlyGlyArgArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67
355 GGAGGTCCGACGCGGTTCGGGGTTCGGCAGCCCTCATCTACTATAGATG 404
67 pAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluGluA 84
405 GGACCTCATGCAGAGGTTGGGAGCCCCAGGAGGATGAGAGAGAGA 454
84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
455 ACGTACAGAGGTTTGGGGGTGATGTGAGACAGCTCATGTGAGAGCTCAGG 504
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHI 117
505 GAAAGGCAGCTGAGCCACAGCCTGCGGGGTTAGCACTACCCGCGCTCA 554
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
555 TCATGACCACCATGATGAGTTTGCCTCATGCCCC 588
```

seq_name: gb_ro:AF051347

```
seq_documentation_block:
LOCUS AF051347 838 bp mRNA ROD 01-OCT-1998
DEFINITION Mus musculus REX-3 mRNA, complete cds.
ACCESSION AF051347
VERSION AF051347.1 GI:3510642
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 838)
Faria,T.N., LaRosa,G.J., Willen,E., Liao,J. and Gudas,L.J.
Characterization of genes which exhibit reduced expression during
the retinoic acid-induced differentiation of F9 teratocarcinoma
cells: involvement of cyclin D3 in RA-mediated growth arrest
Mol. Cell. Endocrinol. 143 (1-2), 155-166 (1998)
99021197
2 (bases 1 to 838)
Faria,T.N., LaRosa,G., Willen,E., Liao,L. and Gudas,L.J.
Direct Submission
Submitted (26-FEB-1998) Pharmacology, Cornell University Medical
College, 1300 York Avenue, New York, NY 10021, USA
JOURNAL Location/Qualifiers
FEATURES
source
1..838
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="F9 teratocarcinoma"
229..579
/note="retinoic acid reduced gene expression"
/codon_start=1
/product="REX-3"
/protein_id="AAC61929.1"
/db_xref="GI:3510643"
/translation="MENDHQKKEEKEKPODTIRREPAVALISEAGKNCAPRGGRRRF
RVROPFIHYRWDLQVRGVEPQGRMRREENVORFGDVRQLMEKLRQLSHSLRAVSTD
PPHHDHDFCLMP"
```

CDS

```
BASE COUNT 238 a 181 c 251 g 168 t
ORIGIN
alignment_scores:
Quality: 688.00 Length: 128
Ratio: 5.375 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

US-09-327-750D-30 x AF051347

```
Align seg 1/1 to: AF051347 from: 1 to: 838
1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspHI 17
|||||
193 ATGAGTCCAAAGATCAAGCGGTGAAAAATCTCAACATGGAGATGACCA 242
17 sGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgArgG 34
|||||
243 TCAGAAAAAGGAGGAGAGAGGAGAAAAAGCCACAAGATACCATCAGAAGG 292
34 luProLaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
|||||
293 AGCCAGCTGTGGCCCTCATCTCCGAGGCTGCCAAAAAATCTGCGCCCTAGA 342
51 GlyGlyArgArgArgPheArgValArgGlnProIleAlaHisTyrArgTr 67
|||||
343 GGAGGTCCGACGCGGTTCGGGGTTCGGCAGCCCTCATCTACTATAGATG 392
67 pAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluGluA 84
|||||
393 GGACCTCATGCAGAGGTTGGGAGCCCCAGGAGGATGAGAGAGAGA 442
84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
|||||
443 ACGTACAGAGGTTTGGGGGTGATGTGAGACAGCTCATGTGAGAGCTCAGG 492
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHI 117
|||||
493 GAAAGGCAGCTGAGCCACAGCCTGCGGGGTTAGCACTACCCGCGCTCA 542
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
543 TCATGACCACCATGATGAGTTTGCCTCATGCCCC 576
```

seq_name: gb_ro:AF097437

```
seq_documentation_block:
LOCUS AF097437 2269 bp DNA ROD 05-AUG-1999
DEFINITION Mus musculus Bex1 protein (Bex1) gene, complete cds.
ACCESSION AF097437
VERSION AF097437.1 GI:5702155
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 2269)
Brown,A.L. and Kay,G.F.
Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
Hum. Mol. Genet. 8 (4), 611-619 (1999)
99172070
MEDLINE 10072429
PUBMED 10072429
REFERENCE 2 (bases 1 to 2269)
Brown,A.L. and Kay,G.F.
AUTHORS Direct Submission
TITLE Submitted (07-OCT-1998) Cancer Unit, Queensland Institute of
JOURNAL Medical Research, Herston Rd, Brisbane, Qld 4030, Australia
FEATURES
source
1..2269
/organism="Mus musculus"
/strain="129SVJ"
/db_xref="taxon:10090"
/chromosome="X"
/map="near Plp"
/tissue_type="liver"
join(409..509,866..942,1263..1873)
/gene="Bex1"
/product="Bex1 protein"
409..1873
/gene="Bex1"
/note="expressed in brain; X-linked"
1268..1654
```

mRNA

gene

CDS

```

/gene="Bex1"
/codon_start=1
/product="Bex1 protein"
/protein_id="RAD47168.1"
/db_xref="GI:5702156"
/translation="MESKDDQGVKNLNMENHDKKEEKPDQITRREPAVALTSEAG
KNCAPRGRRFRVROPTAHYRWDLQMRVGEPOGRMREENVORFGDVRQLMEKLRER
QLSHSLRAVSTDPDHPHDHDFCLMP"
BASE COUNT      563 a   554 c   705 g   447 t
ORIGIN

```

```

alignment_scores:
  Quality: 683.00      Length: 128
  Ratio: 5.378         Gaps: 0
  Percent Similarity: 99.219  Percent Identity: 99.219

```

alignment_block:

US-09-327-750D-30 x AF097437 ..

Align seg 1/1 to: AF097437 from: 1 to: 2269

```

1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspHi 17
|||||
1268 ATGGAGTCCAAAGATCAAGGCGTGAATAATCTCAACATGAGAGTACCA 1317
|||||
17 sGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgArg 34
|||||
1318 TCAGAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1367
|||||
34 luProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
|||||
1368 AGCCAGCTGTGGCCCTGACCTCGAGGTCGCAAAACCTGTGCACCTAGA 1417
|||||
51 GlyClyArgArgArgPheArgValArgGlnProIleAlaHisTyArgTr 67
|||||
1418 GGAGGTGCGAGGGGTTCCGGGTTCCGGCAGCCCTCCTCCTATAGATG 1467
|||||
67 pAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluGluA 84
|||||
1468 GGACCTGATCAGAGGGTTGGGAGGCCCGGAGGAGGATGAGAGGAGA 1517
|||||
84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
|||||
1518 ACGTACAGAGGTTTGGGGTGATGTGAGACAGCTCATGGAGAGCTGAGG 1567
|||||
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||
1568 GAAAGGCAGCTGAGCCACAGCCCTCGGGCGGTAGCACCTGACCCGCTCA 1617
|||||
117 sHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
1618 TCATGACCACCATGATGAGTTTTCCTTCATGCCCC 1651
|||||

```

seq_name: gb_ro:BC003254

```

seq_documentation_block:
LOCUS BC003254      2901 bp      mRNA      ROD      20-FEB-2001
DEFINITION Mus musculus, Similar to dentatorubral pallidolusian atrophy,
clone MGC:5758, mRNA, complete cds.
ACCESSION BC003254
VERSION BC003254.1 GI:13096906
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2901)
AUTHORS Strausberg, R.
DIRECT SUBMISSION
TITLE Submitted (20-FEB-2001) National Institutes of Health, Mammalian
JOURNAL Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

```

REMARK COMMENT

```

USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalon@bcm.tmc.edu.
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 9 Row: 0 Column: 6.

FEATURES Source

```

1..2901
  Location/Qualifiers
    /organism="Mus musculus"
    /db_xref="taxon:10090"
    /clone="MGC:5758 IMAGE:3500522"
    /tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
    ductal carcinoma, 5 month old virgin mouse."
    /clone_lib="NCI-CGAP_Mam6"
    /lab_host="DH10B"
    /note="Vector: pCMV-SPORT6"
    755..2359
    /codon_start=1
    /product="Similar to dentatorubral pallidolusian atrophy"
    /protein_id="AAH03254.1"
    /db_xref="GI:13096907"
    /translation="MVPQFLLTPLTLPGLPKPIHVHTPLHPRAKHEHTPSHSLQSP
    PPQLPPLSPSPAATGPTATQIQEPAEVEYPPSPVPKSPSPKVPVVDVPS
    HASASARENKHLDRGNSCADSLYFVPLEGSKLAKRADLVKVRREAEORAREKE
    REREREREKEREREKERSVKLAQSGRAPVPCSLGVPVHPFPFPGSAVATVPP
    YLGEDTPALRTLSEYARHVASPGNNRHPFVPLGAVDGLLGLYNVALYSSDPAARE
    RERAREDRDLRLKPGPEVKPSELELHGVPGGLDPFPRHGGALQPGPPGLHPFP
    FPHSLPLERLRLAAGPALRPMDSYAEARLAEALGNLQPLARLQMLNVT
    PHHGHSHIHSHLHQDDAIIAASVASVHPLIDPLASGSHLTRIPYPAGTLPNPLPH
    PLHENEVLRHOLFAPYRDLPASLSAPMSAAHQLAQMAHQAELQRLALEQQQWLHAH
    HPLHSVPLPAQEDYXSHLKESDKPL"
BASE COUNT      623 a   958 c   791 g   529 t
ORIGIN

```

alignment_scores:

```

  Quality: 619.00      Length: 118
  Ratio: 5.291         Gaps: 0
  Percent Similarity: 99.153  Percent Identity: 99.153

```

alignment_block:

US-09-327-750D-30 x BC003254

Align seg 1/1 to: BC003254 from: 1 to: 2901

```

1 MetGluSerLysAspGlnGlyValLysAsnLeuAsnMetGluAsnAspHi 17
|||||
223 ATGGAGTCCAAAGATCAAGGCGTGAATAATCTCAACATGAGAGTACCA 272
|||||
17 sGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgArg 34
|||||
273 TCAGAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 322
|||||
34 luProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaProArg 50
|||||
323 AGCCAGCTGTGGCCCTGACCTCGAGGCTGGCAAAACCTGTGCACCTAGA 372
|||||
51 GlyClyArgArgArgPheArgValArgGlnProIleAlaHisTyArgTr 67
|||||

```

373 GGAGGTCCAGCGGTTCCGGGTTCCGACGCCCATCGCTCACTATAGATG 422
67 pAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluGluA 84
|||||
423 GGACCTGATGCAGAGGGTTGGGAGCCCCAGGGAAGGATCAGAGAGAGA 472
84 snValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeuArg 100
|||||
473 ACGTACAGAGGTTTGGGGGTGATGTGACAGAGCTCATGGAGAGCTCAGG 522
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHi 117
|||||
523 GAAAGGAGCTGAGCCACAGCTCGCGGCGTTAGCAGCTACCGCGCTCA 572
117 sHis 118
|||||
573 TCAT 576

seq_name: gb_ro:AF097439

seq_documentation_block:
LOCUS AF097439 785 bp mRNA 13-APR-1999
DEFINITION Mus musculus brain expressed x-linked protein 2 (Bex2) mRNA,
complete cds.
ACCESSION AF097439
VERSION AF097439.1 GI:4580591
KEYWORDS house mouse.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 785)
AUTHORS Brown,A.L. and Kay,G.F.
TITLE Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
JOURNAL Hum. Mol. Genet. 8 (4), 611-619 (1999)
MEDLINE 99172070
REMARK Erratum:([published erratum appears in Hum Mol Genet 1999
May;8(5):943])
REFERENCE 2 (bases 1 to 785)
AUTHORS Brown,A.L. and Kay,G.F.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
FEATURES
Location/Qualifiers
1..785
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="dbEST:w48832"
/chromosome="X"
/map="near Plp"
/dev_stage="embryo; 15.5 dpc"
1..785
/gene="Bex2"
May:8(5):943
139..528
/gene="Bex2"
/codon_start=1
/product="brain expressed x-linked protein 2"
/protein_id="AAD24430.1"
/db_xref="GI:4580592"
/translation="MESKVEQGVKNLNMNDHQEKKEEKPODASKRDPTVALPFEA
GDYVPRGRRFRVROPVIVHWDLHRVGEPOGRMRREENVQFGDDVRLMEKLE
RQLSHSLRAVSTDPPHDDHDFCLMP"

BASE COUNT 212 a 163 c 235 g 175 t
ORIGIN
alignment_scores:
Quality: 589.50 Length: 129
Ratio: 4.912 Gaps: 1
Percent Similarity: 93.023 Percent Identity: 86.822

alignment_block:

US-09-327-750D-30 x AF097439 ..

Align seg 1/1 to: AF097439 from: 1 to: 785

1 MetGluSerLys...AspGlnGlyValLysAsnLeuMetGluAsnAs 16
|||||
139 ATGGAGTCCAAAGTGGAAACAAGGCGTCAAAAATCTCAACATGGAGAATGA 188
16 pHisGlnLysGluGluLysGluGluLysProGlnAspThrIleArgA 33
|||||
189 CCATCAGGAAAGGAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGG 238
33 rgGluProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaPro 49
|||||
239 GGGATCCGATTGTGGCCCTGCCTTTCGAAGCTGGAGACTACTAGTGCCT 288
50 ArgGlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
|||||
289 AGAGGAGGTCCGACGGGTTCCGGGTTCCGACGCCCATCGTCACATACAG 338
66 gTrpAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluG 83
|||||
339 ATGGGACCTCATCATAGGTTGGGAGGCCCGGAGGAAAGGATGAGAGAGG 388
83 luAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeu 99
|||||
389 AGAAGCTACAGAGGTTGGGATGATGTGACAGAGCTCATGGAGAAGCTG 438
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||
439 AGGAAAGGAGCAGCTGAGCCACAGCCCTCGCGGGCGTTAGCACTGACCCGCC 488
116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
489 TCATCATGACCACCATCATGAGTGTTCCTTATGATGATGATGATGATGATG 525

seq_name: gb_pt:AF183416

seq_documentation_block:
LOCUS AF183416 642 bp mRNA 02-SEP-2000
DEFINITION Homo sapiens ovarian granulosa cell 13.0 kDa protein hGR74 homolog
mRNA, complete cds.
ACCESSION AF183416
VERSION AF183416.1 GI:9963770
KEYWORDS human.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 642)
AUTHORS Peng,Y., Qian,B., Yu,Y., Xu,S., Han,Z., Fu,G. and Chen,Z.
TITLE A novel gene expressed in human adrenal gland
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 642)
AUTHORS Peng,Y., GU,W., Huang,C., Xu,S., Han,Z., Fu,G. and Chen,Z.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-1999) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China
FEATURES
Location/Qualifiers
1..642
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="adrenal gland"
3..380
/note="hGR74-h protein"
/codon_start=1
/evidence="not_experimental"
/product="ovarian granulosa cell 13.0 kDa protein hGR74
homolog"
/protein_id="AAG09685.1"
/db_xref="GI:9963771"

TITLE
JOURNAL

Direct Submission
Submitted (21-JAN-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk

COMMENT

On May 13, 1999 this sequence version replaced gi:4582117.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.

This sequence was generated from part of bacterial clone contigs of
human chromosome X, constructed by the Sanger Centre Chromosome X
Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/ChrX>

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at

http://www.sanger.ac.uk/projects/C_elegans/wormpep/RPI-198P4 is
from the library RPCI-1 constructed at the Roswell Park Cancer
Institute by the group of Pieter de Jong. For further details see
<http://bacpac.med.buffalo.edu/>

VECTOR: PCYPAC2

IMPORTANT: This sequence is not the entire insert of clone

RPI-198P4. It may be shorter because we only sequence overlapping
sections once, or longer because we arrange for a small overlap
between neighbouring submissions.

The true left end of clone LLOXNC01-221P2 is at 35615 in this
sequence. The true right end of clone LLOXNC01-73E8 is at 100 in
this sequence.

FEATURES

Location/Qualifiers

1..35714
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q22"
/clone="RPI-198P4"
/clone_lib="RPCI-1"

repeat_region
1..258
/note="L1P3 repeat: matches 5666. 5925 of consensus"
256..537
repeat_region
/note="L1P3 repeat: matches 5391. 5672 of consensus"
1491..1749
repeat_region
/note="AluSc repeat: matches 39. 297 of consensus"
2182..3810
repeat_region
/note="L1P3 repeat: matches 4533. 6150 of consensus"
5169..6687
repeat_region
/note="L1MA2 repeat: matches 4798. 6308 of consensus"
6976..7665
repeat_region
/note="L1MA2 repeat: matches 4113. 4798 of consensus"
8039..9213
repeat_region
/note="L1MA2 repeat: matches 2945. 4113 of consensus"
9523..10437
repeat_region
/note="L1MA2 repeat: matches 2016. 2945 of consensus"
10736..11494
repeat_region
/note="L1MA2 repeat: matches 1262. 2016 of consensus"
11807..12251
repeat_region
/note="L1MA2 repeat: matches 310. 755 of consensus"
12259..12331
repeat_region
/note="HERVFB21 repeat: matches 5. 77 of consensus"
12332..12796
repeat_region
/note="L1R21A repeat: matches 1. 490 of consensus"
16162..16371
repeat_region
/note="105 copies 2 mer tt 55% conserved"
17121..17424
repeat_region
/note="MER33 repeat: matches 1. 322 of consensus"

repeat_region
18259..18290
/note="16 copies 2 mer ca 87% conserved"
18505..18672
repeat_region
/note="MER31A repeat: matches 1. 181 of consensus"
19478..19814
repeat_region
/note="MER2 repeat: matches 1. 345 of consensus"
21061..21101
repeat_region
/note="HERVFB21 repeat: matches 42. 85 of consensus"
21247..21781
repeat_region
/note="MLT1 repeat: matches 44. 568 of consensus"
24465..24492
misc_feature
/note="Single clone region
Single clone region"
24531..24562
misc_feature
/note="Weak data
Weak data"
24563
misc_feature
/note="Single clone region
Single clone region"
24711..24865
misc_feature
/note="region covered by subclones from a PAC DNA PCR
product only at 12x coverage.
region covered by subclones from a PAC DNA PCR product only
at 12x coverage."
24758..25460
misc_feature
/note="CpG island"
/evidence="not_experimental"
24875..24940
repeat_region
/note="33 copies 2 mer cc 66% conserved"
26795..26926
repeat_region
/note="66 copies 2 mer ct 63% conserved"
27088..27133
repeat_region
/note="23 copies 2 mer tg 100% conserved"
27461..27777
repeat_region
/note="L1MB8 repeat: matches 5851. 6171 of consensus"
27802..28232
repeat_region
/note="L1MC/D repeat: matches 5192. 5403 of consensus"
28536..28930
repeat_region
/note="L1MC/D repeat: matches 4802. 5192 of consensus"
29240..29341
repeat_region
/note="TIGGER1 repeat: matches 484. 587 of consensus"
29603..30740
repeat_region
/note="TIGGER1 repeat: matches 587. 1757 of consensus"
31027..31365
repeat_region
/note="TIGGER1 repeat: matches 1757. 2105 of consensus"
31664..31691
repeat_region
/note="TIGGER1 repeat: matches 2105. 2132 of consensus"
31809..32015
repeat_region
/note="TIGGER1 repeat: matches 2132. 2246 of consensus"
32323..32389
repeat_region
/note="TIGGER1 repeat: matches 2246. 2313 of consensus"
34353..34636
repeat_region
/note="MLT1-INTERNAL repeat: matches 607. 889 of
consensus"
35149..35376
repeat_region
/note="MLT1-INTERNAL repeat: matches 1001. 1233 of
consensus"
35523..35670
repeat_region
/note="MLT1 repeat: matches 8. 159 of consensus"
BASE COUNT 9708 a 7094 c 6805 g 12107 t
ORIGIN

alignment_scores:
Quality: 453.50 Length: 129
Ratio: 3.943 Gaps: 3
Percent Similarity: 89.147 Percent Identity: 68.992

alignment_block:
US-09-327-750D-30 x HS198P4/rev ..
Align seg 1/1 to reverse of: HS198P4 from: 1 to: 35714


```

1 MetGluSerLysAsp...GlnGlyValLysAsnLeuAsnMetGluAsnAs 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24358 ATGGAGTCCAAAGAGAAACAGCAGTAACAGTCTCAGCATGGAATATGC 24309

16 pHisGlnLysLysGluGluLysProGlnAspThrIleArgA 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24308 CAACCAAGAAATGAAGAAAGAG.....CAAGTTGCTTAATAAAG 24268

33 rgGluProAlaValAlaLeuLysSerGluAlaGlyLysAsnCysAlaPro 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24267 GGGAGGCC...TTGGCCCTCCCTTTGGATGCTGTGAATACTGTGTGCCT 24221

50 ArgGlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24220 AGAGGAATCGTAGCGGTTCCGCGTTAGCAGCCCATCTCGCAGTATAG 24171

66 gTrpAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluG 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24170 ATGGGATATGATCATAGCTTGGAGAACCCACAGGCAAGGATGAGAGAAG 24121

83 luAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeu 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24120 AGAATATGGAAGATGGGAGGAGGTGAGCAGCTGATGGAAAGCTG 24071

100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24070 AGGGAAGCAGTTGAGTCATAGCTCGCGGCAGTCAGCACTGACCCCC 24021

116 oHisHisAspHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24020 TCACCATGACCATCATGATGAGTTTGGCTTATGCC 23984

```

seq_name: gb_pr:AF237783

```

seq_documentation_block: 791 bp mRNA PRI 02-SEP-2000
LOCUS AF237783
DEFINITION Homo sapiens brain-expressed protein BEX1 (BEX1) mRNA, complete
          cds.
ACCESSION AF237783
VERSION AF237783.1 GI:9963898
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 791)
AUTHORS Yang,Q.S., Ying,K., Xie,Y. and Mao,Y.M.
TITLE A Novel Human X-linked Brain Expressed Protein
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 791)
AUTHORS Mao,Y.M., Xie,Y., Yang,Q.S., Wu,H., Lin,S. and Ying,K.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2000) Genetic Research Institute, Fudan
          University, 220 Handan Rd., Shanghai 200433, P.R.China
FEATURES
          Location/Qualifiers
            source
              1..791
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="X"
                /map="between DXS990 and DX1059 by RH (Stanford G3)"
            gene
              1..791
            CDS
              172..549
                /gene="BEX1"
                /codon_start=1
                /product="brain-expressed protein BEX1"
                /protein_id="AAG09752.1"
                /db_xref="GI:9963898"
                /translation="MESKEKLAVNSLSEMANQENKEQVANKGEPLALPLDAGEYCV
                VPGNRRFRVROPILQYRWDMHRLGEPQNRWENRIGEGVQRLMEKLRKQLS
                HSLRAVSTDPHPHHDHDFCLMP"

```

BASE COUNT
ORIGIN

214 a 172 c 218 g 187 t

```

alignment_scores:
  Quality: 450.50
  Ratio: 3.952
  Percent Similarity: 88.372
  Length: 129
  Gaps: 3
  Percent Identity: 68.992

alignment_block:
US-09-327-750D-30 x AF237783
Align seg 1/1 to: AF237783 from: 1 to: 791

1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
172 ATGGAGTCCAAAGAGAAACAGCAGTAACAGTCTCAGCATGGAATATGC 221

16 pHisGlnLysLysGluGluLysProGlnAspThrIleArgA 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
222 CAACCAAGAAATGAAGAAAGAG.....CAAGTTGCTTAATAAAG 262

33 rgGluProAlaValAlaLeuLysSerGluAlaGlyLysAsnCysAlaPro 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
263 GGGAGGCC...TTGGCCCTCCCTTTGGATGCTGTGAATACTGTGTGCCT 309

50 ArgGlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
310 AGAGGAATCGTAGCGGTTCCGCGTTAGCAGCCCATCTCGCAGTATAG 359

66 gTrpAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluG 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
360 ATGGGATATGATCATAGCTTGGAGAACCCACAGGCAAGGATGAGAGAAG 409

83 luAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeu 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
410 AGAATATGGAAGATGGGAGGAGGTGAGCAGCTGATGGAAAGCTG 459

100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
460 AGGGAAGCAGTTGAGTCATAGCTCGCGGCAGTCAGCACTGACCCCC 509

116 oHisHisAspHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
510 TCACCATGACCATCATGATGAGTTTGGCTTATGCC 546

seq_name: gb_pr:AF251053

seq_documentation_block:
LOCUS AF251053
DEFINITION Homo sapiens X-linked protein mRNA, complete cds.
ACCESSION AF251053
VERSION AF251053.1 GI:13625167
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 744)
AUTHORS Mao,Y., Xie,Y., Zhou,Z., Zhao,W., Wang,W., Huang,Y.,
          Wang,S., Tang,R., Chen,X. and Wu,C.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2000) Institute of Genetics, School of Life
          Sciences, Fudan University, 220 Handan Road, Shanghai 200433, P.R.
          China
FEATURES
          Location/Qualifiers
            source
              1..744
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
            CDS
              101..487
                /codon_start=1
                /product="X-linked protein"
                /protein_id="AAK34943.1"
                /db_xref="GI:13625168"
                /translation="MESKEERLNNLIVENQENDEKDEQVANKGEPLALPLNVS

```

BASE COUNT 222 a 134 c 208 g 180 t
ORIGIN
EYCVPRGRRRRRVPQPILOYRWDIMHRLGEPQARMREENMERIGEEVRLMEKLREK
QLSHSLRAVSTDPPIHHHDEFCLMP"

alignment_scores:
Quality: 450.00 Length: 129
Ratio: 3.879 Gaps: 2
Percent Similarity: 89.922 Percent Identity: 66.667

alignment_block:

US-09-327-750D-30 x AF251053

Align seg 1/1 to: AF251053 from: 1 to: 744

```
1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
|||||
101 ATGGAGTCCAAAGAGGACGAGCGTTAAACAATCTCATCTGGAATAATGT 150
16 pHisGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgA 33
|||||
151 CAACAGGAAATGATGAAAAGATGAAAAGGAGCAAGTTGCTAATAAAG 200
33 rgGluProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaPro 49
|||||
201 GGGAGCCC...TTGGCCCTACCTTTGAATCTTAGTGAATACTGTGTCCT 247
50 ArgGlyClyArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
|||||
248 AGAGGAACCGTAGGCGGTTCCGCGTTAGGAGCCCATCTCGAGTATAG 297
66 gTTPAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluG 83
|||||
298 ATGGACATATATCATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGAG 347
83 luAsnValGlnArgPheGlyClyAspValArgGlnLeuMetGluLysLeu 99
|||||
348 AGAATATGGAAGAGATTGGGAGAGGTGAGACAGCTGATGGAAGAGCTG 397
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||
398 AGGGAAGACAGTGTAGTCTAGTCTCGCGGCGAGTCAGCACTGATCCCC 447
116 oHisHisAspHisAspGluPheCysLeuMetPro 128
|||||
448 TCACCATGACCATCAGATGAGTTTGGCTTATGCCCC 484
```

seq_name: gb_pat:AX078272

seq_documentation_block:
LOCUS AX078272 898 bp DNA PAT 22-FEB-2001
DEFINITION Sequence 76 from Patent WO0107471.
ACCESSION AX078272
VERSION AX078272.1 GI:13157963
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 898)
TITLE Hillman,J.L., Lal,P., Tang,Y.T., Yue,H., Au-Young,J., Bandman,O.,
Azimzai,Y., Yang,J., Lu,D.A., Baughn,M.R., Patterson,C. and Shah,P.
JOURNAL Cell cycle and proliferation proteins
Patent: WO 0107471-A 76 01-FEB-2001;
Incyte Genomics, Inc. (US)

FEATURES
source
1..898
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 5664154CB1"
BASE COUNT 250 a 186 c 251 g 211 t
ORIGIN

alignment_scores:

Quality: 450.00 Length: 129
Ratio: 3.879 Gaps: 2
Percent Similarity: 89.922 Percent Identity: 66.667

alignment_block:

US-09-327-750D-30 x AX078272

Align seg 1/1 to: AX078272 from: 1 to: 898

```
1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
|||||
254 ATGGAGTCCAAAGAGGACGAGCGTTAAACAATCTCATCTGGAATAATGT 303
16 pHisGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgA 33
|||||
304 CAACAGGAAATGATGAAAAGATGAAAAGGAGCAAGTTGCTAATAAAG 353
33 rgGluProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaPro 49
|||||
354 GGGAGCCC...TTGGCCCTACCTTTGAATCTTAGTGAATACTGTGTCCT 400
50 ArgGlyClyArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
|||||
401 AGAGGAACCGTAGGCGGTTCCGCGTTAGGAGCCCATCTCGAGTATAG 450
66 gTTPAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluG 83
|||||
451 ATGGACATATATCATAGGCTTGGAGAGCCACAGGCAAGGATGAGAGAG 500
83 luAsnValGlnArgPheGlyClyAspValArgGlnLeuMetGluLysLeu 99
|||||
501 AGAATATGGAAGAGATTGGGAGAGGTGAGACAGCTGATGGAAGAGCTG 550
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||
551 AGGGAAGACAGTGTAGTCTAGTCTCGCGGCGAGTCAGCACTGATCCCC 600
116 oHisHisAspHisAspGluPheCysLeuMetPro 128
|||||
601 TCACCATGACCATCAGATGAGTTTGGCTTATGCCCC 637
```

seq_name: gb_pr:HSV870H8

seq_documentation_block:

LOCUS HSV870H8 31321 bp DNA PRI 23-NOV-1999
DEFINITION Human DNA sequence from cosmid V870H8, between markers DXS366 and DXS87 on chromosome X contains ESTs.

ACCESSION Z70233
VERSION Z70233.1 GI:1235542
KEYWORDS X.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 31321)
TITLE Direct Submission
JOURNAL Submitted (19-MAR-1995) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, UK. E-mail enquires: humquery@sanger.ac.uk

COMMENT
IMPORTANT: This sequence is not the entire insert of clone V870H8. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.

The true left end of clone V870H8 is at 1 in this sequence. The true left end of clone V693A8 is at 3121.

V870H8 is from the human chromosome X-specific cosmid library.

Location/Qualifiers

1..31321
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

/chromosome="x"
/map="x"
/clone="GHC-870H8"
/clone_lib="SCcv"
475..701
/note="L1 element fragment"
repeat_region
802..988
/note="L1 element fragment"
repeat_region
2533..4138
/misc_feature
/note="match: multiple ESTs"
4189..4246
/note="29 copies of 2 mer 91 % conserved"
repeat_region
7496..7874
/note="L1 element fragment"
repeat_region
8022..8108
/note="L1 element fragment"
repeat_region
8302..8493
/note="L1 element fragment"
repeat_region
8585..8671
/note="L1 element fragment"
repeat_region
8769..9010
/note="L1 element fragment"
repeat_region
9259..9384
/note="L1 element fragment"
repeat_region
9624..9731
/note="L1 element fragment"
repeat_region
9781..9948
/note="L1 element fragment"
repeat_region
9890..9948
/note="L1 element fragment"
repeat_region
10418..10592
/note="L1 element fragment"
repeat_region
10593..10688
/note="Alu repeat: matches 298..108 of consensus"
repeat_region
10698..10781
/note="Alu repeat: matches 96..1 of consensus"
repeat_region
11112..11403
/note="L1 element fragment"
repeat_region
11922..12004
/note="Alu repeat: matches 1..308 of consensus"
repeat_region
12062..12308
/note="L1 element fragment"
repeat_region
12423..12707
/note="L1 element fragment"
repeat_region
12710..12916
/note="Alu repeat: matches 302..1 of consensus"
repeat_region
12979..13081
/note="L1 element fragment"
repeat_region
13473..13532
/note="L1 element fragment"
repeat_region
15871..15928
/note="L1 element fragment"
repeat_region
18180..18417
/note="L1 element fragment"
repeat_region
18778..19056
/note="Alu repeat: matches 1..308 of consensus"
repeat_region
22279..22350
/note="L1 element fragment"
repeat_region
22449..22995
/note="L1 element fragment"
repeat_region
23239..23374
/note="L1 element fragment"
repeat_region
23371..23606
/note="L1 element fragment"
repeat_region
23395..23541
/note="L1 element fragment"
repeat_region
23705..23778
/note="L1 element fragment"
repeat_region
23805..23960
/note="L1 element fragment"

```

```

repeat_region 24011..24541
/note="L1 element fragment"
repeat_region 24638..24770
/partial
/note="Alu repeat: matches 150..1 of consensus"
24829..30975
/note="L1 element fragment"
BASE COUNT 9462 a 6147 c 6129 g 9583 t
ORIGIN
alignment_scores:
Quality: 450.00 Length: 129
Ratio: 3.879 Gaps: 2
Percent Similarity: 89.922 Percent Identity: 66.667
alignment_block:
US-09-327-750D-30 x HSV870H8
Align seg 1/1 to: HSV870H8 from: 1 to: 31321
1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
|||||
3508 ATGGAGTCCAAAGAGGAGCAAGCGTTAAACATCTCATCGTGGAAAATGT 3557
16 pHisGlnLysGluGluLysGluGluLysProGlnAspThrIleArgA 33
|||||
3558 CAACAGGAGAAATGATGAAAAGATGAAAAGAGGAGCAAGTTGCTAATAAAG 3607
33 rGluProAlaValAlaLeuIleSerGluAlaGlyLysAsnCysAlaPro 49
|||||
3608 GGGAGGCC...TTGGCCCTACCTTTGAATGTTAGTGAATACTGTGTGCCT 3654
50 ArgGlyClyArgArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
|||||
3655 AGAGGAACCGTAGGCGGTTCCCGCTTAGCGAGCCCATCTCGAGTATAG 3704
66 gTTPAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluG 83
|||||
3705 ATGGACATATATCATAGGCTTGGAGAGCCAGGCAAGGATGAGAGAGG 3754
83 luAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeu 99
|||||
3755 AGAATATGGAAGGATTGGGAGGAGGTGAGACAGCTGATGGAAGAGCTG 3804
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
|||||
3805 AGGAAAGACAGTTGAGTCATAGTCTCGGGCAGTCAGCAGCTGATCCCC 3854
116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
3855 TCACCATGACCATCAGCATGAGTTTGCCTTATGCC 3891
seq_name: gb_pr:AL133348
seq_documentation_block:
LOCUS AL133348 40584 bp DNA PRI 06-MAR-2000
DEFINITION Human DNA sequence from clone RPL-79P11 on chromosome Xq21.32-22.1.
Contains the gene for a novel protein similar to mouse Bex2
(brain-expressed X-linked protein 2), ESTs, STSs, GSSs and a
putative CpG island, complete sequence.
ACCESSION AL133348
VERSION AL133348.8 GI:7076386
KEYWORDS HTG; Bex2; CpG island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 40584)
AUTHORS Wilson,S.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone

```

COMMENT

requests: clonerequest@sanger.ac.uk
 On Feb 24, 2000 this sequence version replaced gi:6997869.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence has been finished according to sequence map criteria
 as follows. An attempt is made to resolve all sequencing problems,
 such as compressions and repeats, but not necessarily within known
 annotated human repeat sequence elements (e.g. Alu). Where the
 sequence is ambiguous, there is an annotation using the 'unsure'
 feature key.
 The following abbreviations are used to associate primary accession
 numbers given in the feature table with their source databases:
 Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
 on the WORMPEP database can be found at
<http://www.sanger.ac.uk/projects/C.elegans/wormpep> This sequence
 was generated from part of bacterial clone contigs of human
 chromosome X, constructed by the Sanger Centre Chromosome X Mapping
 Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/ChrX>
 RPI-79P11 is from the library RPCI-1 constructed at the Roswell
 Park Cancer Institute by the group of Pieter de Jong. For further
 details see <http://bacpac.med.buffalo.edu/>
 VECTOR: pCYPAC2
 IMPORTANT: This sequence is not the entire insert of clone
 RPI-79P11 it may be shorter because we only sequence overlapping
 sections once, or longer because we arrange for a small overlap
 between neighbouring submissions.
 The true left end of clone LLOXNC01-105G4 is at 40485 in this
 sequence. The true right end of clone LLOXNC01-177E8 is at 100 in
 this sequence.

FEATURES

Source	Location/Qualifiers
repeat_region	1..40584
repeat_region	/organism="Homo sapiens"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="X"
repeat_region	/map="q21.32-22.1"
repeat_region	/clone="RPI-79P11"
repeat_region	/clone_lib="RPCI-1"
repeat_region	19..292
repeat_region	/note="AluJo repeat: matches 3..275 of consensus"
repeat_region	470..1284
repeat_region	/note="LIMB5 repeat: matches 5318..6157 of consensus"
repeat_region	1295..1898
repeat_region	/note="L1PA15 repeat: matches 5556..6157 of consensus"
repeat_region	1899..2345
repeat_region	/note="L1PA15 repeat: matches 4922..5370 of consensus"
repeat_region	2323..2698
repeat_region	/note="LIMB5 repeat: matches 4960..5343 of consensus"
repeat_region	2927..3065
repeat_region	/note="MER74A repeat: matches 164..305 of consensus"
repeat_region	3655..9795
repeat_region	/note="L1PA4 repeat: matches 2..6142 of consensus"
repeat_region	9866..9982
repeat_region	/note="FLAW_C repeat: matches 6..124 of consensus"
repeat_region	9989..10254
repeat_region	/note="L1 repeat: matches 4027..4293 of consensus"
repeat_region	10253..10852
repeat_region	/note="L1M4 repeat: matches 4755..5333 of consensus"
repeat_region	10853..11260
repeat_region	/note="MLT1B repeat: matches 3..390 of consensus"
repeat_region	11261..11474
repeat_region	/note="L1M4 repeat: matches 4538..4755 of consensus"
repeat_region	11480..11563
repeat_region	/note="WSTB repeat: matches 1..80 of consensus"
repeat_region	11564..12475
repeat_region	/note="L1 repeat: matches 3039..3971 of consensus"
misc_feature	14463..14861
repeat_region	/note="match: GSS: Em:AQ729215"
repeat_region	15578..15854
misc_feature	Complement(16007..16556)
repeat_region	/note="match: GSS: Em:AQ772810"
repeat_region	16197..16569
repeat_region	/note="L1M4 repeat: matches 4267..4627 of consensus"
repeat_region	16629..16714
repeat_region	/note="L1M4C repeat: matches 1578..1666 of consensus"
repeat_region	16902..17133
repeat_region	/note="L1M4 repeat: matches 0..242 of consensus"
repeat_region	17536..17728
repeat_region	/note="MIR repeat: matches 1..191 of consensus"
repeat_region	17746..18323
repeat_region	/note="L2 repeat: matches 1599..2175 of consensus"
repeat_region	18706..18755
repeat_region	/note="L1MB3 repeat: matches 6034..6083 of consensus"
repeat_region	18756..19236
repeat_region	/note="MER66B repeat: matches 1..486 of consensus"
repeat_region	19237..21108
repeat_region	/note="MER66-internal repeat: matches 4808..6676 of consensus"
repeat_region	21109..21163
repeat_region	/note="MLT2B repeat: matches 391..445 of consensus"
repeat_region	21164..21555
repeat_region	/note="MER66B repeat: matches 3..391 of consensus"
repeat_region	21557..21662
repeat_region	/note="L1MB3 repeat: matches 6080..6185 of consensus"
repeat_region	21685..21761
repeat_region	/note="L1M repeat: matches 5299..5392 of consensus"
repeat_region	21762..21929
repeat_region	/note="L1MB8 repeat: matches 5902..6063 of consensus"
repeat_region	21930..22212
repeat_region	/note="AluSg repeat: matches 3..284 of consensus"
repeat_region	22213..22321
repeat_region	/note="L1MB8 repeat: matches 6063..6173 of consensus"
repeat_region	22333..22740
repeat_region	/note="L1MD2 repeat: matches 5867..6288 of consensus"
repeat_region	22744..23094
repeat_region	/note="L1MD repeat: matches 2..342 of consensus"
repeat_region	23218..23524
repeat_region	/note="AluSg repeat: matches 1..305 of consensus"
repeat_region	23603..23947
repeat_region	/note="L1M4 repeat: matches 4916..5184 of consensus"
repeat_region	23948..24218
repeat_region	/note="AluJo repeat: matches 1..282 of consensus"
repeat_region	24219..24510
repeat_region	/note="L1M4 repeat: matches 4601..4916 of consensus"
repeat_region	24508..24621
repeat_region	/note="L1M4 repeat: matches 4374..4488 of consensus"
repeat_region	24650..25077
repeat_region	/note="MLT1C repeat: matches 1..466 of consensus"
repeat_region	25078..25914
repeat_region	/note="L1M4 repeat: matches 3497..4346 of consensus"
repeat_region	25920..26741
repeat_region	/note="L1M4 repeat: matches 2202..3065 of consensus"
repeat_region	26762..26906
repeat_region	/note="L1PA7 repeat: matches 6001..6145 of consensus"
repeat_region	26907..27135
repeat_region	/note="L1PA5 repeat: matches 5674..5902 of consensus"
repeat_region	27266..27509
repeat_region	/note="L1M4 repeat: matches 2164..2080 of consensus"
repeat_region	27557..27769
repeat_region	/note="L1MEC repeat: matches 253..460 of consensus"
repeat_region	28120..28167
repeat_region	/note="L2 repeat: matches 2703..2750 of consensus"
repeat_region	29043..29349
repeat_region	/note="MER58B repeat: matches 1..341 of consensus"
repeat_region	30389..30452
repeat_region	/note="16 copies 4 mer tata 89% conserved"
repeat_region	Complement(30502)
repeat_region	/gene="dJ79P11.1"
repeat_region	Complement(join(30502..31137,31514..31589,32016..>32134))
repeat_region	/gene="dJ79P11.1"
repeat_region	/note="match: ESTs: Em:AA317587 Em:AA358632 Em:AI214048

```
Em:AI929470 Em:AA319216 Em:AA317625 Em:D81392 Em:AA317418
Em:AA365528 Em:C15186 Em:AI291270 Em:AA160660 Em:AA612605
Em:AW160832 Em:AA384405 Em:AA160657 Em:AA081744
Em:AI291126 Em:C15591 Em:AA384268 Em:AI929703 Em:AI302026
Em:W19547 Em:AI141727 Em:AA946933 Em:AI141583 Em:H61106
Em:W60581 Em:AI826617 Em:AA706545 Em:AA464771 Em:AA862327
Em:W68380 Em:AA641652 Em:AA384269 Em:AA527696 Em:W48778
Em:AA633989 Em:AI659486 Em:AW167301 Em:AI929428
Em:AA637941 Em:C15948 Em:AI745669 Em:AW024587 Em:AW379207;
match: proteins: Tr:O9Y516 Tr:O9Y517
/evidence=not_experimental
/product="dj79p11.1 (novel protein similar to mouse Bex2
complement(30502, .32134)
/gene="dj79p11.1"
30507..30988
/notes="match: STS: Em:G24641; match: STS: Em:T86927"
30508..30852
/notes="match: STS: Em:R43117; match: STS: Em:G23964"
complement(30518, .30523)
/gene="dj79p11.1"
30541..30922
/notes="match: STS: Em:H61107"
31543..32262
/notes="CpG island"
/evidence=not_experimental
31647..31730
/notes="42 copies 2 mer cc 64% conserved"
complement(32655, .33316)
/notes="match: GSS: Em:AA0489303"
33027..34514
/notes="L1ME2 repeat: matches 4631..6155 of consensus"
34528..34786
/notes="L1ME3A repeat: matches 4404..4664 of consensus"
34787..35085
/notes="Alus9 repeat: matches 1..300 of consensus"
35086..35412

alignment_scores:
  Quality: 450.00      Length: 129
  Ratio: 3.879        Gaps: 2
Percent Similarity: 89.922 Percent Identity: 66.667

alignment_block:
US-09-327-750D-30 x AL133348/rev ..
Align seg 1/1 to reverse of: AL133348 from: 1 to: 40584

1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
|||||
31132 ATGGAGTCCAAAGAGGAGGAGCGTTAAACAATCTCATCTGCGAAAATGT 31083

16 pHisGlnLysLysGluGluLysGluLysProGlnAspThrIleArgA 33
|||||
31082 CAACACGAGAAATGATGAAAGAGATGAAAGGACGAGTGTCTAATAAG 31033

33 rgGluProAlaValAlaLeuSerGluAlaGlyLysAsnCysAlaPro 49
|||||
31032 GGGAGGCC...TTGGCCCTACCTTTGAATGTTAGTGAATACGTGTGCCT 30986

50 ArgGlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
|||||
30985 AGAGGAAACCGTAGGCGGTTCCGCGTTAGCAGCCCATCTCGCAGTATAG 30936

66 gTrpAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGlu 83
|||||
30935 ATGGGACATAATCATAGCTTGGAGAGCCACAGGCAAGGATGAGAGG 30886

83 luAsnValGlnArgPheGlyGlyAspValArgGlnLeuMetGluLysLeu 99
|||||
30885 AGAATATGGAAAGGATTTGGGAGGAGGTGAGACAGCTGATGAGAAAGCTG 30836

100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
```

```
|||||:::|||||
30835 AGGAAAGACAGTTGAGTCATAGTTTTCGGGCGAGTCAGCACTGATCCCC 30786

116 oHisHisAspHisHisAspGluPheCysLeuMetPro 128
|||||
30785 TCACCATGACCATCAGCATGAGTTTTCGGCTATATGCC 30749

seq_name: gb_htg:HSU80B1

seq_documentation_block:
LOCUS HSU80B1 41029 bp DNA HTG 10-JUL-2001
DEFINITION Homo sapiens chromosome X clone LLOXNC01-80B1 map q22.1-22.3, ***
SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION AL022169
VERSION AL022169.3 GI:13276704
KEYWORDS HTG; HTGS_PHASE2; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 41029)
AUTHORS Bird,C.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Mar 12, 2001 this sequence version replaced gi:12750927.
COMMENT
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: cu80B1
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator ABI; 19% of reads
Chemistry: Dye-terminator Big Dye; 80% of reads
Consensus quality: 41009 bases at least Q40
Consensus quality: 41023 bases at least Q30
Consensus quality: 41027 bases at least Q20
Insert size: 41029; sum-of-contigs
Insert size: 51751; 0.6% error; agarose-fp
Quality coverage: 9.67x in Q20 bases; sum-of-contigs Quality
coverage: 7.67x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
      source
      1..41029
         /organism="Homo sapiens"
         /db_xref="taxon:9606"
         /chromosome="X"
         /map="q22.1-22.3"
         /clone="LLOXNC01-80B1"
         /clone_lib="LLOXNC01"
      1..41029
      misc_feature
         /note="assembly_fragment:00890"
BASE COUNT 13249 a 7796 c 7781 g 12203 t
ORIGIN

alignment_scores:
  Quality: 450.00      Length: 129
  Ratio: 3.879        Gaps: 2
Percent Similarity: 89.922 Percent Identity: 66.667

alignment_block:
US-09-327-750D-30 x HSU80B1
```

Align seq 1/1 to: HSU80B1 from: 1 to: 41029

1 MetGluSerLysAspGln...GlyValLysAsnLeuAsnMetGluAsnAs 16
18412 ATGGAGTCCAAAGAGGAACGCGTTAAACATCTCATCGTGGAAATGT 18461
16 pHisGlnLysLysGluGluLysGluGluLysProGlnAspThrIleArgA 33
18462 CAACGAGAAATCATGAAAGAGATGAAAGGACCAAGTTGCTAATAAAG 18511
33 rgGluProAlaValAlaLeuSerGluAlaGlyLysAsnCysAlaPro 49
18512 GGGAGCCC...TTGGCCCTACTCTTGATGTAGTGAATACTGTGTGCCT 18558
50 ArgGlyGlyArgArgPheArgValArgGlnProIleAlaHisTyrAr 66
18559 AGAGNAACCGTAGCGGTTCCGCGTTAGCGACGCCATCTCGCATATAG 18608
66 gTrpAspLeuMetGlnArgValGlyGluProGlnGlyArgMetArgGluG 83
18609 ATGGACATAATGCATAGGCTTGACGAGCCACAGCAAGGATGAGAGG 18658
83 luAsnValGlnArgPheGlyClyAspValargGlnLeuMetGluLysLeu 99
18659 AGAATATGGAAGGATTGGGAGGAGGTGACAGAGCTGATGGAAAGACTG 18708
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProPr 116
18709 AGGCAAAAGCAGTTGACTCATAGTCTCGGGCAGTCAGCACTGATCCCC 18758
116 oHisHisAspHisAspGluPheCysLeuMetPro 128
18759 TCACCATCACCATCAGCATGAGTTTGCCCTTATGCC 18795

seq_name: gb_sts:G24641

seq_documentation_block:	
LOCUS	G24641 504 bp DNA STS 31-MAY-1996
DEFINITION	human STS WI-11354, sequence tagged site.
ACCESSION	G24641
VERSION	G24641.1 GI:1344967
KEYWORDS	STS; STS sequence; primer; sequence tagged site.
SOURCE	human STSS derived from sequences in dbEST and the Unigene collection.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 504)
AUTHORS	Hudson,T.
TITLE	Whitehead Institute/MIT Center for Genome Research; Physically Mapped STSS
JOURNAL	Unpublished (1995)

Primer A: TTTTGTGCTTTTAAAGGTGC
Primer B: TTACGTGGTCTCCTATTACCA
STS size: 150
PCR Profile:

PCR profile:
 Presoak:
 Denaturation:
 Annealing: 56 degrees C
 Polymerization:
 PCR Cycles: 35
 Thermal Cycler:

```

Primer: each 5 pM
dNTPs: each 4 mM
Taq Polymerase: 0.025 units/ul
Total Vol: 20 ul

Buffer:
MgCl2: 1.5 mM
KCl: 50 mM
Tris-HCL: 10 mM
pH: 9.3

Derived from dbEST (genbank accession T86927).
FEATURES             Location/Qualifiers
     source            1..504
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
     STS               1..150
     primer_bind       1..22
     primer_bind       complement(129..150)
     BASE COUNT       134 a 118 c 101 g 144 t 7 others
     ORIGIN

alignment_scores:
    Quality: 299.00      Length: 85
    Ratio: 4.153        Gaps: 3
    Percent Similarity: 84.706  Percent Identity: 71.765

alignment_block:
US-09-327-750D-30 x G24641/rev ..

Align seg 1/1 to reverse of: G24641 from: 1 to: 504

47  CysAlaProArgGly.GlyArgArgArgPhe.ArgValArgGlnProfile 62
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
503 TGTCCCCCTANAGGAAANCCGTAGCGCGTTCCCGCGTTAGGCANCCCCAC 65

63 AlaHisTyrArgTrpAspLeuMetGlnArgValGlyGlu.ProGlnGlyA 79
    ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::
453 CGCAGTANAGATGGGATATGATGCATAGGCTTGAGAGAACCCACAGGCAA 70

79 rgMetArgGluGluAsnValGlnArgPheGlyGlyAspValArgGlnLeu 95
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
403 GGATGAGAGACAGAAATATGGAAGAGATTGGGGAGAGGTGAGACAGCTG 35

96 MetGluLysLeuArgGluuArgGlnLeuSerHisLeuArgAlaValise 111
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
353 ATGGAAAAGCTGAGGAAAACGACGTTGAGTCANAGCTCTCGCGGCAGTCAG 30

112 rThrAspProProHisHisAspHisHisAspGluPheCysLeuMetPro 128
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
303 CACTGACCCCTCACCATGACCATCATGATGAGTTTTCGTTTATGCC 255

```

OM of: US-09-327-750D-34 to: EST:* out_format : pfs
Date: Mar 11, 2002 2:16 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=Cgnt2_1/USPTO.spool/US09327750/runat_11032002_101153_20308/app_query.fasta_1.1472
-DB=EST -OPMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0
-ALIGN=15 -MODE=LOCAL -OUTPWT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09327750_@CGNI_1_5654
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-34

Query length: 118

Database: EST:*

Database sequences: 11351937

Database length: 1077921985

Search time (sec): 4085.940000

score_list:

Sequence	Strd	Orig	zScore	EScore	Len	Documentation
gb_est1:AI227867	-	620.00	1031.44	2.7e-48	568	! AI227867 EST224562 Normalized r
gb_est2:BI289546	-	592.00	984.02	1.2e-45	643	! BI289546 UI-R-DKO-cff-c-12-0-UI
gb_hic:AK010400	+	546.00	906.12	2.6e-41	789	! AK010400 Mus musculus ES cells
gb_est2:BG083261	+	546.00	906.10	2.6e-41	792	! BG083261 H3086C08-5 NIA Mouse 1
gb_est1:AA473525	-	534.00	885.85	4.1e-40	458	! AA473525 vg78a09.r1 Barstead MP
gb_est1:AAW251884	-	520.00	865.45	4.8e-39	540	! AAW251884 UI-R-Bj0-adn-b-06-0-UI
gb_est2:BG070341	-	510.00	845.61	5.4e-38	747	! BG070341 H3086C08-3 NIA Mouse 1
gb_est2:BF723075	+	503.00	837.44	1.8e-37	519	! BF723075 mb27e12.y1 Soares_NME
gb_est1:BE333774	+	483.00	804.95	1.1e-35	463	! BE333774 us27h06.y1 Soares_NME
gb_est2:W17712	+	466.00	779.24	3.1e-34	317	! W17712 mb77e04.r1 Soares mouse
gb_est1:AA117429	+	444.00	742.75	3.3e-32	312	! AA117429 mb27e12.r1 Beddington
gb_est1:AI012535	-	426.00	710.42	2.1e-30	444	! AI012535 EST206986 Normalized r
gb_est2:W35893	+	425.00	708.34	2.7e-30	472	! W35893 mc53g09.r1 Soares mouse
gb_est2:W54487	+	410.00	683.49	6.6e-29	465	! W54487 md09c11.r1 Soares mouse
gb_est1:AL118340	+	407.00	679.16	1.2e-28	422	! AL118340 v91112b26 Beddington mc
gb_est2:BF722027	-	406.00	680.14	1.0e-28	286	! BF722027 mb27e12.x1 Soares_NME
gb_est1:AA104150	+	399.00	665.68	6.5e-28	433	! AA104150 mc50h03.r1 Life Tech m
gb_est2:W85403	+	388.00	647.24	6.9e-27	442	! W85403 mf58d02.r1 Soares mouse
gb_est1:AA893925	-	386.00	644.71	9.5e-27	393	! AA893925 EST197728 Normalized r
gb_est1:AA956955	-	385.00	641.88	1.4e-26	467	! AA956955 UI-R-E1-f1-f-07-0-UI.S
gb_est2:BF556502	+	373.00	623.91	1.4e-25	348	! BF556502 UI-R-E1-f1-f-07-0-UI.X
gb_est1:AL517301	+	372.00	618.13	2.9e-25	639	! AL517301 AL517301 LTI_NFL011.NE
gb_est2:BE883414	+	372.00	617.30	3.2e-25	722	! BE883414 601509167f1 NIH.MGC.71
gb_est1:AL544124	+	372.00	616.23	3.7e-25	845	! AL544124 AL544124 LTI_NFL006.PI
gb_est1:AU117899	+	372.00	616.08	3.8e-25	865	! AU117899 AU117899 HEMBA1 Homo S
gb_est2:BG122197	+	372.00	616.07	3.8e-25	866	! BG122197 602347948f1 NIH.MGC.90
gb_est1:AL545120	+	372.00	615.77	3.9e-25	905	! AL545120 AL545120 LTI_NFL006.PI
gb_est2:BG709427	+	372.00	615.59	4.0e-25	929	! BG709427 602673481f1 NIH.MGC.96
gb_est1:BG035675	+	372.00	615.26	4.2e-25	975	! BG035675 602325745f1 NIH.MGC.90
gb_est1:AL560650	+	372.00	615.21	4.2e-25	983	! AL560650 AL560650 LTI_NFL010.BC
gb_est1:AL539004	+	372.00	614.94	4.3e-25	1023	! AL539004 AL539004 LTI_FL013.FE
gb_est1:AL549875	+	372.00	614.80	4.4e-25	1044	! AL549875 AL549875 LTI_NFL006.F
gb_est2:BF930215	+	367.50	611.31	6.9e-25	579	! BF930215 IL5-NT0277-111200-319-
gb_est1:AL534630	+	367.00	610.64	7.5e-25	566	! AL534630 AL534630 LTI_FL013.FBI
gb_est1:AL523320	+	365.50	603.99	1.7e-24	1043	! AL523320 AL523320 LTI_NFL003.N
gb_est1:AW385267	+	364.00	606.04	1.9e-24	534	! AW385267 RC0-LT0001-261199-011-
gb_est2:BG527431	+	364.00	603.54	1.9e-24	772	! BG527431 602557227f1 NIH.MGC.59
gb_est2:BG009563	-	362.50	602.97	2.0e-24	581	! BG009563 QV1-GN0319-021200-526-
gb_est1:AL573904	+	362.00	597.32	4.2e-24	1182	! AL573904 AL573904 LTI_NFL006.F
gb_est2:BI040887	-	360.50	600.82	2.7e-24	489	! BI040887 QV3-NT0278-120201-502-

gb_est1:AL582060 - 359.50 594.05 6.3e-24 1037 ! AL582060 AL582060 LTI_NFL01
gb_est2:BI040889 - 356.50 593.18 7.1e-24 565 ! BI040889 QV3-NT0278-120201-5
gb_est1:AW934883 + 355.00 589.55 1.1e-23 668 ! AW934883 RCL-DT0001-230200-0
gb_est1:AW934905 + 355.00 589.38 1.2e-23 685 ! AW934905 RCL-DT0001-230200-0
gb_est1:AW934769 + 354.50 588.78 1.2e-23 662 ! AW934769 RCL-DT0001-230200-0
seq_name: gb_est1:AI227867
seq_documentation_block:
LOCUS AI227867 568 bp mRNA EST 20-JAN-1999
DEFINITION EST224562 Normalized rat brain, Bento Soares Rattus sp. cDNA clone
RBCRM82 3' end, mRNA sequence.
ACCESSION AI227867
VERSION AI227867.1 GI:3811754
KEYWORDS EST.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 568)
AUTHORS Lee N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,
Kerlavager, A.R. and Adams, M.D.
TITLE Rat Genome Project: Generation of a Rat EST (EST) Catalog & Rat
Gene Index
JOURNAL Unpublished (1998)
COMMENT Other ESTs: TC58086
Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.
Location/Qualifiers
1..568
/organism="Rattus sp."
/db_xref="ATCC (inhost):2036891"
/db_xref="taxon:10118"
/clone="RBCRM82"
/clone_lib="Normalized rat brain, Bento Soares"
/note="Organ: brain; Vector: pT73pac; Site:1: EcoRI;
Site 2: NotI"
BASE COUNT 140 a 143 c 108 g 177 t
ORIGIN
alignment_scores:
Quality: 620.00 Length: 118
Ratio: 5.299 Gaps: 0
Percent Similarity: 99.153 Percent Identity: 99.153
alignment_block:
US-09-327-750D-34 x AI227867/rev ..
Align seg 1/1 to reverse of: AI227867 from: 1 to: 568
1 MetAlaSerLysValLysGlnValLleLeuAspLeuThrValGluLysAs 17
|||||
528 ATGGCGTCCAAAGTCAAAAGTCATCTCGATCTCCTGCGAAGA 479
|||||
17 pLysLysAsnLysLysGlyLysAlaSerLysGlnSerGluGluGlu 34
|||||
478 CAAAGAAAACAAAAGTGGGAGGCGCTCCAAACAAAGTGAAGAAGAT 429
|||||
34 erHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 50
|||||
428 CCACCATCTGGAAGAGGTGTAACAAAGAGCTGGGGCAATGTCAGG 379
|||||
51 ArgLysValArgArgLeuValProAsnPhLeuTrpAlaIleProAsnAr 67
|||||
378 AGAAAGTCAAGCGACTTGTGCCTAACTTCTATGGCCATACCTAATAG 329
|||||

67 gHisValAspHisSerGluGlyGlyGluValGlyArgPheValGlyG 84
|||||
328 GCATGTTGATCACAGTGAAGGGGAGAGGTTGGGAGATTCGTAGGGC 279
|||||
84 InValMetGluAlaLysArgHisSerLysGluGlnGlnMetArgProTyr 100
|||||
278 AGGTGATGGAAGCCAGAGAAAGTCTAAGGAGCAACAGATGAGGCCCTTAC 229
|||||
101 ThrArgPheArgThrProGluProAspAsnHisTyrAspPheCysLeuI 117
|||||
228 ACGGTTTCCGAACCCCTGAACCTGACAATATTACGACTTTTGCCCTCAT 179
|||||
117 ePro 118
|||||
178 ACCT 175

seq_name: gb_est2:BI289546

seq_documentation_block:
LOCUS BI289546 643 bp mRNA EST 19-JUL-2001
DEFINITION UI-R-DK0-cff-c-12-0-UI.s1 UI-R-DK0 Rattus norvegicus cDNA clone
UI-R-DK0-cff-c-12-0-UI 3', mRNA sequence.
ACCESSION BI289546
VERSION BI289546
KEYWORDS
SOURCE Norway rat.
Rattus norvegicus
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 643)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized rat kidney pool library cDNA Library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 1-51,
>AT-rich#Low_complexity
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
Source Location/Qualifiers
1..643
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DK0-cff-c-12-0-UI"
/clone_lib="UI-R-DK0"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; The UI-R-DK0
library is a subtracted library derived from a mixture of
five individually tagged normalized rat libraries:
brain-nRBP (20%), heart-nRBP (20%), kidney-nRBP (20%),
aorta-nRAP (20%), and placenta-nRBP (20%). Each original
library was constructed from a mixture of equal amounts of
RNA from seven different developmental time-points:
embryonic day 17, embryonic day 19, embryonic day 21,

adult day 1, adult day 12, adult day 75, and adult day
200. (Exception: the aorta pool does not contain embryonic
day 17 RNA and the placenta pool contains only the three
embryonic stages). Each library was normalized
individually according to the procedure described by
Bonaldo, Lennon & Soares (Genome Research Genome 6:
791-806, 1996). For construction of the DK0 subtracted
library, plasmid DNA from each of the five individually
tagged normalized libraries was mixed in the proportions
specified above and electroporated into competent bacteria
for production of single-stranded circular DNA
representing the pool of libraries. Single-stranded
circular DNA representing these five normalized libraries
was then used as a tracer in a subtractive hybridization
with a driver (PCR amplified inserts from a plasmid DNA
template preparation) comprising: a) a set of about 1,000
arrayed clones from each of the five non-normalized
libraries of brain (CTOs), heart (CSOs), kidney (CUOs),
aorta (CWOs), and placenta (CXOs). The resulting pool of
approximately 5,000 clones represented about 33.3% of the
final driver population. A set of about 2,000 arrayed
clones from each of the five normalized libraries of brain
(CTO), heart (CSO), kidney (CUO), aorta (CWO), and
placenta (CXO). The resulting pool of about 10,000 clones
represented about 66.6% of the final driver population.

TAG-LIB=UI-R-DK0

TAG-TISSUE=rat kidney pool

TAG-SEQ=CAAGACTGTCTC

BASE COUNT 160 a 149 c 113 g 221 t
ORIGIN

alignment_scores:

Quality: 592.00 Length: 118
Ratio: 5.060 Gaps: 0
Percent Similarity: 99.153 Percent Identity: 98.305

alignment_block:

US-09-327-750D-34 x BI289546/rev ..

Align seg 1/1 to reverse of: BI289546 from: 1 to: 643

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
|||||
601 ATGCGGTCCAAAGTCAACAAAGTCATCTGATCTCCTCTGGGAGAAAGA 552
|||||
17 pLysLysAsnLysLysGlyLysAlaSerLysGlnSerGluGluGlu 34
|||||
551 CAAGAAAACAAAAAGGTGGGAG.GCCTCCAAACAAAGTGAAGAGAAT 503
|||||
34 erHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 50
|||||
502 CCCACCATCTGGGAAGAGGTTGAAACAAAGAGCCTGGGGCAATGTCCAG 453
|||||
51 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 67
|||||
452 AGGAAAGTCAAGCGCACTGTGCTTAACCTTTCTATGGGCCCATCTAATAG 403
|||||
67 gHisValAspHisSerGlyGlyGlyGluGluValGlyArgPheValGlyG 84
|||||
402 GCATTTGATCATCAGTGAAGGGGAGAGGAGGTTGGGAGATTCGTAGGGC 353
|||||
84 InValMetGluAlaLysArgHisSerLysGluGlnGlnMetArgProTyr 100
|||||
352 AGGTGATGGAAGCCAGAGAAAGTCTAAGGAGCAACAGATGAGGCCCTTAC 303
|||||
101 ThrArgPheArgThrProGluProAspAsnHisTyrAspPheCysLeuI 117
|||||
302 ACGGTTTCCGAACCCCTGAACCTGACAATATTACGACTTTTGCCCTCAT 253
|||||
117 ePro 118
|||||
252 ACCT 249

seq_name: gb_hlc:AK010400

seq_documentation_block:

LOCUS AK010400 789 bp mRNA HTC 05-JUL-2001
DEFINITION Mus musculus ES cells cDNA, RIKEN full-length enriched library, clone:2410004M13, full insert sequence.

ACCESSION AK010400

VERSION AK010400.1 GI:12845816

KEYWORDS CAP trapper.

SOURCE Mus musculus (strain:C57BL/6J) ES cells cDNA to mRNA.

clone_lib:RIKEN full-length enriched mouse cDNA library

clone:2410004M13.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS 1 (bases 1 to 789)

TITLE Carninci, P. and Hayashizaki, Y.

JOURNAL High-efficiency full-length cDNA cloning

METHODS in enzymology. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE

AUTHORS 2 (bases 1 to 789)

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome research. 10 (10), 1617-1630 (2000)

JOURNAL 20499374

MEDLINE 11042159

REFERENCE

AUTHORS 3 (bases 1 to 789)

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multipillar sequencer

Genome research. 10 (11), 1757-1771 (2000)

JOURNAL 20530913

MEDLINE 11076861

REFERENCE

AUTHORS 4 (bases 1 to 789)

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

JOURNAL 5 (bases 1 to 789)

PUBMED 11076861

REFERENCE

AUTHORS 5 (bases 1 to 789)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Soabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,

Fax: 81-45-503-9216)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for

further details.

CDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in RIKEN

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

COMMENT

Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAAGGATCCAGAGCTCTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCGAGTTAATTAAATCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

Host: SOLR.

Location/Qualifiers

1..789

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/db_xref="MGI:1328321"

/db_xref="MGI:1908964"

/clone="2410004M13"

/cell_type="ES cells"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

208..564

/note="putative"

/codon_start=1

/protein_id="BAB26911.1"

/db_xref="GI:12845817"

/translation="MASRFKOVILDLTVEKDKKGGKASKQSEEPHLEEVENKK

PGSNVRRKVRRLVNFNLWALPNRHRNREGGDDVGRFVVQGVTEVKKRKTTEQQVRPYRR

FRTPEDNHYDFCLIP"

770..775

/note="putative"

789

/note="putative"

BASE COUNT 239 a 161 c 205 g 184 t

ORIGIN

alignment_scores:

Quality: 546.00 Length: 118

Ratio: 4.964 Gaps: 0

Percent Similarity: 93.220 Percent Identity: 87.288

alignment_block:

US-09-327-750D-34 x AK010400 ..

Align seg 1/1 to: AK010400 from: 1 to: 789

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17

|||||

208 ATGGCATCCAAATTTAAACAAGTCATACTGGATCTCACTGGGAGAAGA 257

|||||

17 pLysLysAsnLysLysGlyLysAlaSerLysGlnSerGluGluGluS 34

|||||

258 CAAAAAAGCAAAAAAGGTGGGAGGCTCCAAACAAAGTGAAGAAGAAC 307

|||||

34 erHisHisLeuGluValGluAsnLysLysProGlyGlyAsnValArg 50

|||||

308 CCCACCATCTGGAGAGGTGGAACACAGAGGCTGGGGAAATGTCGA 357

|||||

51 ArgLysValArgArgLeuValProAsnPhleLeuTrpAlaIleProAsn 67

|||||

358 AGCAAGATCAGGCGACTTGTGCCTAACTTCTCTGGCCCATACCAATAG 407

|||||

67 gHisValAspHisSerGlyGlyGluGluValGlyArgPheValGlyG 84

|||||

408 GCATGTTGATCGCAATCAAGGGGAGAGGATGTTGGAGATTTGTAGTC 457

|||||

84 lnValMetGluAlaLysArgHisSerLysGluGlnGlnMetArgProTyr 100

|||||

458 AGGACACAGAGTCACAGAGAAAGACTACGAGGAGGAGGTCAGGCTTAC 507

|||||

101 ThrArgPheArgThrProGluProAspAsnHisTyrAspPheCysLeu 117

|||||

508 AGCGTTTCGGAACCCGGAACCTGACAATCATTACGACTTTTGCCTCAT 557

117 epro 118

|||||
558 ACCT 561

seq_name: gb_est2:BG083261

seq_documentation_block: 792 bp mRNA EST 26-JAN-2001
LOCUS BG083261 H3086C08-5 NIA Mouse 15K CDNA Clone Set Mus musculus CDNA clone
DEFINITION H3086C08 5', mRNA sequence.
ACCESSION BG083261
VERSION BG083261.1 GI:12565829
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 792)
AUTHORS Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
J.T.S., Carter,M.G. and Ko,M.S.H.
TITLE Verification and initial annotation of NIA mouse 15K CDNA clone set
JOURNAL Unpublished (2001)
COMMENT Other ESTs: H3086C08-3
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://lgsun.grc.nia.nih.gov/cdna/15k.html> for details.
Plate: H3086 row: C column: 08
Seq primer: -21M13 Reverse
High quality sequence stop: 792
POLYA-No.

FEATURES

Location/Qualifiers
1..792
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="H3086C08"
/clone.lib="NIA Mouse 15K CDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT) -Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."
247 a 153 c 210 g 179 t 3 others

BASE COUNT
ORIGIN

alignment_scores:
Quality: 546.00
Ratio: 4.964
Length: 118
Gaps: 0

Percent Similarity: 93.220 Percent Identity: 87.288

alignment_block:
US-09-327-750D-34 x BG083261

Align seg 1/1 to: BG083261 from: 1 to: 792

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
|||||
190 ATGGCATCCAAATTTAAACAAGTCATACCTGATCTCCTGTTGGAGAAAGA 239
|||||
17 pLysLysAsnLysLysGlyLysAlaSerLysGlnSerGluGluGlu 34
|||||
240 CAAAAAGACAAAAAGGTGGGAAGCCTCCAAACAAGTGAAGAAGAAC 289
|||||
34 erHisLysLeuGluValGluAsnLysLysProGlyGlyAsnValArg 50
|||||
290 CCCACCATCTGGAAGAGTTGAAACAAGAAGCTGGGGAAATGTCGA 339
|||||
51 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 67
|||||
340 AGAAAGTCAGGCGACTTGTGCCTAACTTCTCTGGGCCATACCAATAG 389
|||||
67 gHisValAspHisSerGluGlyGluValGluValArgPheValClyG 84
|||||
390 GCATGTTGATCCCAATGAAGGGGAGAGGATGTTGGAGCATTTGTAGTGC 439
|||||
84 lnValMetGluAlaLysArgHisSerLysGluGlnGlnMetArgProTyr 100
|||||
440 AGGGAACAGAAAGTCAAGACAAAGACTACGAGCAGCAGGTAGGCGCTTAC 489
|||||
101 ThrArgPheArgThrProGluProAspAsnHisTyrAspPheCysLeuII 117
|||||
490 AGCGGTTTCGGAACCCGGAACCTGACAATCATTACGACTTTTGCCTCAT 539
|||||
117 epro 118
|||||
540 ACCT 543

seq_name: gb_est1:AA473525

seq_documentation_block:

LOCUS AA473525 458 bp mRNA EST 18-JUN-1997
DEFINITION vg78a09.r1 Barstead MPLRBL Mus musculus. CDNA clone IMAGE:872056 5' similar to SW:HG74_HUMAN Q00994 OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74. ?, mRNA sequence.

ACCESSION AA473525

VERSION AA473525.1 GI:2201752

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 458)

REFERENCE

AUTHORS

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and

Waterston,R.

TITLE

The WashU-HMHI Mouse EST Project

Unpublished (1996)

JOURNAL

Comment

Contact: Marra M/Mouse EST Project

WashU-HMHI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:511536

Seq primer: -28m13 rev2 ET from Amersham

High quality sequence stop: 417.

FEATURES
source

Location/Qualifiers
1..458
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="IMAGE:872056"
/clone_lib="Barstead MPLRB1"
/sex="mixed"
/tissue_type="Kidney"
/dev_stage="6 weeks"
/lab_host="DH10B"
/note="vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TCTTACCAATCTGAAGTGGGAGCGCCCTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[CATGGATCGGTACCl, digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead."

BASE COUNT 147 a 88 c 115 g 108 t

ORIGIN

alignment_scores:
Quality: 534.00 Length: 116
Ratio: 4.944 Gaps: 0
Percent Similarity: 93.103 Percent Identity: 86.207

alignment_block:

US-09-327-750D-34 x AA473525

Align seg 1/1 to: AA473525 from: 1 to: 458

3 SerLysValLysGlnValIleLeuAspLeuThrValLysAspLysLy 19
:::|||||
13 GCCAAATTAACAAGTACATCTGATCTACCTGTGGAGAAACAAAA 62
19 sAsnLysLysGlyGlyAlaSerLysGlnSerGluGluSerHis 36
:::|||||
63 AGACAAAAAGGTGGAGGCGCTCCAAACAAAGTGAAGAACCCACC 112
36 istLeuGluValGluAsnLysLysProGlyGlyAsnValArgLys 52
113 ATCTGAAGAGGTTGAAACAAAGAGCCTGGGGAAATGTCGAAGAA 162
53 ValArgLeuValProAsnPheLeuTTPAlaIleProAsnArgHisVa 69
163 GTCAGGCGACTTGCTTCTCTGCGCCATACCAATAGGCATGT 212
69 lAspHisSerGluGlyGlyGluValGlyArgPheValGlyGlnVal 86
213 TGATCGCAATGAAGGGGAGAGGATGTTGGAGATTGTAGTCAGGAA 262
86 etGluAlaLysArgHisSerLysGluGlnMetArgProTyrThrArg 102
263 CAGAAGTCAAGAGAGACTACGAGCAGCAGGTGAGGCTTACAGCGGT 312
103 PheArgThrProGluProAsnHisTyrAspPheCysLeuIlePro 118
313 TTCGGAACCCCGAAGCTGACAAATCATACGACTTTTGCTCATACCT 360

seq_name: gb_est1:AW251884

seq_documentation_block:
LOCUS AW251884 540 bp mRNA EST 17-DEC-1999
DEFINITION UI-R-BJ0-adn-b-06-0-UI.s1 UI-R-BJ0 Rattus norvegicus cDNA clone
UI-R-BJ0-adn-b-06-0-UI 3', mRNA sequence.
ACCESSION AW251884
VERSION AW251884.1 GI:6595475
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 540)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.wesg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized atrium at 16.5 dpc library cDNA Library Preparation:
M.B. Soares Lab Clone distribution: clones will be available
through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source

Location/Qualifiers
1..540
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BJ0-adn-b-06-0-UI"
/clone_lib="UI-R-BJ0"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-BJ0
library is a subtracted library derived from the UI-R-AAL,
UI-R-AB1, UI-R-AC1, UI-R-AD1, UI-R-AE1, UI-R-AF1, and
UI-R-AG1 libraries. These libraries represent tissues from
rat atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal
at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV
canal at 15 dpc, and ventricle at 13 dpc. The tag is a
string of 5-6 nucleotides present between the Not I site
and the oligo-dT track. The library was constructed as
described by Bonaldo, Lennon and Soares, Genome Research
6: 791-806, 1996.
TAG_LIB=UI-R-BJ0
TAG_TISSUE=atrium at 16.5 dpc
TAG_SEQ=GATTC"

BASE COUNT 145 a 120 c 94 g 180 t

ORIGIN

alignment_scores:
Quality: 520.00 Length: 97
Ratio: 5.417 Gaps: 0
Percent Similarity: 98.969 Percent Identity: 98.969

alignment_block:

US-09-327-750D-34 x AW251884/rev

Align seg 1/1 to reverse of: AW251884 from: 1 to: 540

22 LysGlyGlyLysAlaSerLysGlnSerGluGluSerHisLeuG1 38
538 AAAAGTGGGAAGCCCTCCAAACAAAGTGAAGAGGAATCCACCATCTGGA 489
38 uGluValGluAsnLysLysProGlyGlyAsnValArgArgLysValArg 55
488 AGAGGTGAAACAAAGAGCCTGGGGCAATGTGAGGAGGAAGATCAGGC 439
55 rGluValProAsnPheLeuTrpAlaIleProAsnArgHisValAspHis 71

438 GACTTGTGCGCTAACTTTCTATGCGCCATACCTAATAGGCATGTTGATCAC 389
 72 SerGluGlyGlyGluGluValGlyArgPheValcylGlnValMetGluAl 88
 |||||
 388 AGTGAAGGGGAGAGAGAGTTGGGAGATTCCTAGGCAGGTGATGGAGC 339
 88 aLysArgHisSerLysGluGlnGlnMetArgProTyrThrArgPheArgT 105
 |||||
 338 CAAGAGAAAGCTCTAAGGAGCAACAGATGAGGCCTTACACGCGTTCCGAA 289
 105 hrProGluProAspAsnHisTyrAspPheCysLeuIlePro 118
 288 CCCCTGAACCTGACAAATCATTAGGACTTTTGCCTCATACCT 248

seq_name: gb_est2:BG070341

seq_documentation_block: 747 bp mRNA EST 26-JAN-2001
 LOCUS BG070341
 DEFINITION H3086C08-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
 H3086C08 3', mRNA sequence.
 ACCESSION BG070341
 VERSION BG070341.1 GI:12552910
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 747)
 Kargul, J.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka
 T.S., Carter, M.G. and Ko, M.S.H.
 Verification and initial annotation of NIA mouse 15K cDNA clone set
 Unpublished (2001)
 Other ESTs: H3086C08-5
 Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please
 visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.
 Plate: H3086 row: C column: 08
 Seq primer: -21M13 Forward
 High quality sequence stop: 747
 POLYA=Yes.

FEATURES
 Source

Location/Qualifiers
 1..747
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="H3086C08"
 /clone.lib="NIA Mouse 15K cDNA Clone Set"
 /sex="Clones arrayed from a variety of cDNA libraries"
 /dev_stage="Clones arrayed from a variety of cDNA libraries"
 /lab_host="DH10B"
 /note="Vector: pSPOR1; Site 1: SalI; Site 2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT) primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum

BASE COUNT 178 a 198 c 140 g 231 t
 ORIGIN
 alignment_scores:
 Quality: 510.00 Length: 118
 Ratio: 4.722 Gaps: 0
 Percent Similarity: 91.525 Percent Identity: 85.593

alignment_block:

us-09-327-750D-34 x BG070341/rev ..

Align seg 1/1 to reverse of: BG070341 from: 1 to: 747

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
 |||||
 581 ATGGCATCCAAATTTAAACAAGTCATAGTGGATCTCAGTGTGGAGAAGA 532
 17 pLysLysAsnLysLysGlyLysAlaSerLysGlnSerGluGluGluS 34
 |||||
 531 CAAAAAGACAAAAAGGTGGGAGGCTCCAAACAAAGTGAAGAGAAC 482
 34 erHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 50
 |||||
 481 CCCACCATCTGGAGAGGTGTAACAAGAGCCTGGGGGAATGTCCGA 432
 51 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 67
 |||||
 431 AGGAAAGTCAGGCGACTTGTGCCTAACTTTCTCTGGGCCATACCAATAG 382
 67 gHisValAspHisSerGluGlyGlyGluGluValGlyArgPheValcylG 84
 |||||
 381 GCATGTTGATCGCAATGAAGGGGAGAGGATTTGGGAGATTGTGTAGTGC 332
 84 InValMetGluAlaLysArgHisSerLysGluGlnGlnMetArgProTyr 100
 |||||
 331 AGGGAACAGAGTCACAGAGAAAGACTACGCGAGCAGCGTGAGCGT.TAC 283
 101 ThrArgPheArgThrProGluProAspAsnHisTyrAspPheCysLeuI 117
 |||||
 282 AGGCGTTTCCGAACCCCGGAACCTGACATCATTAGGACTTTTGCCTCAT 233
 117 ePro 118
 |||||
 232 ACCT 229

seq_name: gb_est2:BF723075

seq_documentation_block:

LOCUS BF723075 519 bp mRNA EST 03-JAN-2001
 DEFINITION mab27e12.y1 Soares_NMEBA_branchial_arch Mus musculus cDNA clone
 IMAGE:3971447 5' similar to TR:Q9Y516 Q9Y516 DJ635G19.2.3 ;, mRNA
 sequence.
 ACCESSION BF723075
 VERSION BF723075.1 GI:12024077
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 519)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Other ESTs: mab27e12.x1
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
 , Ph.D.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Washington University Genome Sequencing Center

JOURNAL

COMMENT

/clone="IMAGE:538606"
 /clone_lib="Beddington mouse embryonic region"
 /sex="pooled"
 /tissue_type="embryo"
 /dev_stage="7.5dpc"
 /lab_host="DH12S"

/note="Organ: whole embryo; Vector: pCMV-SPORT; Site_1:
 SalI; Site_2: NotI; Cloned unidirectionally. Primer:
 Oligo dt. Gastrulating embryos were collected at 7.5dpc
 from C57Bl6 x DBA matings, excluding embryos that had
 developed head folds and all extraembryonic tissues.
 Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).
 Referenced in Development 121, 2479-2489 (1995)"

BASE COUNT 110 a 56 c 92 g 54 t
 ORIGIN

alignment_scores:
 Quality: 444.00 Length: 100
 Ratio: 4.774 Gaps: 0
 Percent Similarity: 93.000 Percent Identity: 86.000

alignment_block:
 US-09-327-750D-34 x AAL17429 ..

Align seg 1/1 to: AAL17429 from: 1 to: 312

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
 12 ATGGCATCCAAATTTAAACAAGTCATACCTGGATCTCCTGAGGAAAGA 61
 17 pLysLysAsnLysGlyGlyLysAlaSerLysGlnSerGluGluGlu 34
 62 CAAAAAAGACAAAAGGTGGGAAGGCTCCAAACAAAGTGAAGAGAAC 111
 34 exHisLeuGluGluValGluValGluAsnLysLysProGlyGlyAsnValArg 50
 112 CCCACCATCTGGAAGAGGTGAAACAAAGAGCCTGGGGGAAATGTCGCA 161
 51 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 67
 162 AGGAAAGTCAGCGACTGTGCTTACTTCTCTGGGCCATACCTAATAG 211
 67 gHisValAspHisSerGluGlyGlyGluValGlyArgPheValGlyG 84
 212 GCATGTTGATCCATGAGAGGGGAGAGGATGTTGGGAGATTTGTAGTCC 261
 84 InValMetGluAlaLysArgHisSerLysGluGlnGlnMetArgProTy 100
 262 AGGGAACAGAGTCAAGAGAAAGACTACGGAGCAGCAGGTGAGGCTTAC 311

seq_name: gb_est1:AI012535

seq_documentation_block:
 LOCUS AI012535 444 bp mRNA EST 15-JUN-1998
 DEFINITION EST206986 Normalized rat placenta, Bento Soares Rattus sp. CDNA
 clone RPLAY54 3' end, mRNA sequence.

ACCESSION AI012535
 VERSION AI012535.1 GI:3226367
 KEYWORDS EST.
 SOURCE Rattus sp.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 444)

AUTHORS Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,
 Kerlavage, A.R. and Adams, M.D.

TITLE Rat Genome Project: Generation of a Rat EST (EST) Catalog & Rat
 Gene Index

JOURNAL Unpublished (1998)

COMMENT Contact: Lee, NH
 The Institute for Genomic Research

9712, Medical Center Drive, Rockville, MD 20850, USA
 Tel: (301)-838-3529
 Fax: (301)-838-0208
 Email: nhlee@igr.org
 Seq primer: M13-21.

FEATURES

source

1..444
 /Organism="Rattus sp."
 /db_xref="taxon:10118"
 /clone="RPLAY54"
 /clone_lib="Normalized rat placenta, Bento Soares"
 /note="Organ: placenta; Vector: p7T3pac; Site_1: EcoRI;
 Site_2: NotI"

BASE COUNT 135 a 104 c 81 g 124 t
 ORIGIN

alignment_scores:
 Quality: 426.00 Length: 80
 Ratio: 5.462 Gaps: 0
 Percent Similarity: 97.500 Percent Identity: 97.500

alignment_block:
 US-09-327-750D-34 x AI012535/rev ..

Align seg 1/1 to reverse of: AI012535 from: 1 to: 444

39 GluValGluAsnLysLysProGlyGlyAsnValArgArgLysValArgAr 55
 444 GAGGTGAAACATGAAGCCTGGGGCAATGTCAGGAGGAAAGTCAGGCG 395
 55 gLeuValProAsnPheLeuTrpAlaIleProAsnArgHisValAspHis 72
 394 ACTTGTGCTTACTTCTATGGCCATACCTAATAGCATGTTGATCACA 345
 72 erGluGlyGlyGluValGlyArgPheValGlyGlnValMetGluAla 88
 344 GTGAAGGGGAGAGGAGGTGGAGATTCTGTAGGCGAGGTGATGGAAGCC 295
 89 LysArgHisSerLysGluGlnGlnMetArgProTyThrArgPheArgTh 105
 294 AAGAAAGTCTAAGGAGCAACAGATGAGCGCTTACACGCTTCCGAAC 245
 105 rProGluProAspAsnHisTyAspPheCysLeuIlePro 118
 244 CCTGAACCTGACATCATACGACTTTTGCCTCATACCT 205

seq_name: gb_est2:W35893

seq_documentation_block:
 LOCUS W35893 472 bp mRNA EST 14-MAY-1996
 DEFINITION mc53909.r1 Soares mouse embryo NME13.5 14.5 Mus musculus CDNA
 clone IMAGE:352288 5' similar to SW:HG74_HUMAN Q00994 OVARIAN
 GRANULOSA CELL 13.0 KD PROTEIN HGR74. [1] ;, mRNA sequence.

ACCESSION W35893

VERSION W35893.1 GI:1317733

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lennon, G., Soares, B., Willson, R. and
 Waterston, R.

TITLE The WashU-HMIT Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project

WashU-HMIT Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:224088
 Seq primer: ETPprimer
 High quality sequence stop: 441.
 Location/Qualifiers
 1..472
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone_lib="IMAGE:352288"
 /clone_lib="Soares mouse embryo NbME13.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTCACCAATCTGAAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
 State Univ., from 2]; double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT7T3 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M.Fatima Bonaldo."
 BASE COUNT 149 a 95 c 157 g 71 t
 ORIGIN

alignment_scores:
 Quality: 425.00 Length: 97
 Ratio: 4.722 Gaps: 0
 Percent Similarity: 92.784 Percent Identity: 85.567
 alignment_block:
 US-09-327-750d-34 x W35893 ..
 Align seg 1/1 to: W35893 from: 1 to: 472
 1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
 |||||
 181 ATGGCATCCAAATTTAAACAAGTCATCTGGATCTCCTGCTGGAGAAAGA 230
 17 pLysLysAsnLysLysGlyGlyLysAlaSerLysGlnSerGluGluGlu 34
 |||||
 231 CAAAGAAACAAAGAGTGGGAGGCTCCAAACAAGAGTGAAGAGAAC 280
 34 erHishisLeuGluGluValGluAsnLysLysProGlyGlyAsnValArg 50
 |||||
 281 CCCACCATCTGGAAGAGTGTGAACAAGAGCTGGGGGAAATGTCGGA 330
 51 ArgLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 67
 |||||
 331 AGGAAAGTCAGGCGACTTGTGCGCTAACTTCTCTGGGCCATACCAATAG 380
 67 gHisValAspHisSerGluGlyGlyGluGluValGlyArgPheValGlyG 84
 |||||
 381 GCATGTTGATCGCAATGAGGGGACAGGATGTTGGAGATTTGTAGTCG 430
 84 lnValMetGluAlaLysArgHisSerLysGluGlnGlnMet 97
 |||||
 431 AGGGAACAGAGTCAAGAGAGAAAGACTACGAGCAGCAGGCTG 471
 seq_name: gb_est2.W54487
 seq_documentation_block:
 LOCUS W54487 465 bp mRNA EST 03-JUN-1996

DEFINITION md09c11.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
 clone IMAGE:367892 5', mRNA sequence.
 ACCESSION W54487
 VERSION W54487.1 GI:1355547
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 REFERENCE 1 (bases 1 to 465)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilton,R. and
 Waterston,R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:229324
 Seq primer: ETPprimer
 High quality sequence stop: 328.
 Location/Qualifiers
 1..465
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:367892"
 /clone_lib="Soares mouse embryo NbME13.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTCACCAATCTGAAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
 State Univ., from 2]; double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT7T3 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M.Fatima Bonaldo."
 BASE COUNT 147 a 90 c 156 g 72 t
 ORIGIN
 alignment_scores:
 Quality: 410.00 Length: 99
 Ratio: 4.505 Gaps: 0
 Percent Similarity: 91.919 Percent Identity: 84.848
 alignment_block:
 US-09-327-750d-34 x W54487 ..
 Align seg 1/1 to: W54487 from: 1 to: 465
 1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
 |||||
 168 ATGGCATCCAAATTTAAACAAGTCATCTGGATCTCCTGCTGGAGAAAGA 217
 17 pLysLysAsnLysLysGlyGlyLysAlaSerLysGlnSerGluGluGlu 34
 |||||
 218 CAAAGAAACAAAGAGTGGGAGGCTCCAAACAAGAGTGAAGAGAAC 267

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on:	March 11, 2002, 15:42:51 ; Search time 144 Seconds (without alignments) 4167.553 Million cell updates/sec
---------	---

Title: US-09-327-750D-28

Perfect score:

Sequence: 1 acgagcgctctggccagcagc.....ggggcttqtqtqccagtqa 700

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0.

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length:	2000000000
Maximum DB seq length:	2000000000

Post-processing: Minimum Match 0%

Fast processing: Minimum Match 0%
Maximum Match 100%

Maximum match 100%
Listing first 45 summaries

Database : N Geneseq 1101:*

1: /SID2/cgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID2/cgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID2/cgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID2/cgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID2/cgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID2/cgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID2/cgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID2/cgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID2/cgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID2/cgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID2/cgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID2/cgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID2/cgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID2/cgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID2/cgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID2/cgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID2/cgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID2/cgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID2/cgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID2/cgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID2/cgdata/geneseq/geneseq/NA2000.DAT.*
22: /SID2/cgdata/geneseq/geneseq/NA2001.DAT.*
23: /SID2/cgdata/geneseq/geneseq/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	693.6	99.1	700	22	AAE233528	Mouse NADE DNA. M
2	307.4	43.9	891	22	AAE233529	Human NADE DNA. H
3	284.6	40.7	917	21	AAE21748	Human breast and o
4	253.6	36.2	532	21	AAAC01005	Human secreted pro
5	192.2	27.5	187	21	AAAC01004	Human secreted pro
6	147.2	21.0	187	21	AAAE87147	Rat hepatocyte car
7	127.2	18.2	426	22	AAH823586	Human Ovarian tumo
8	80.2	11.5	865	22	AAH03517	Human cDNA clone (
9	80.2	11.5	1229	22	AAH13750	Human cDNA sequenc
10	80.2	11.5	1364	22	AAE85548	cDNA encoding cD1F
11	58	8.3	401	21	AAAC01003	Human secreted pro

12	56.4	8.1	792	22	AAH45143	Human brain expres
13	55.8	8.0	767	22	AAH75810	Human X chromosome
14	55.8	8.0	898	22	AAF59611	Human cell cycle a
15	55.6	7.9	662	21	AAC03880	Human secreted pro
16	55.6	7.9	692	21	AAC10889	Human secreted pro
17	55.6	7.9	698	21	AAC03879	Human secreted pro
18	55.6	7.9	858	22	AAI60367	Human polynucleoti
19	54.2	7.7	862	22	AAI58581	Human polynucleoti
20	50.4	7.2	298	21	AAC06182	Human secreted pro
21	44.6	6.4	5552	21	AAA58309	pBIP/PyCSP.1 plasm
22	42.2	6.0	128	21	AAC12100	Human secreted pro
23	41.4	5.9	830	22	AAI22045	Probe #11978 for g
24	41.4	5.9	830	22	AAI47337	Probe #16023 used
25	41.4	5.9	830	22	AAI07740	Probe #7731 used t
26	41	5.9	279	21	AAC06181	Human secreted pro
27	40.8	5.8	381	22	AAI12579	Probe #2512 for ge
28	40.8	5.8	381	22	AAI33926	Probe #2612 used t
29	40.8	5.8	381	22	AAI02485	Probe #2476 used t
30	40.4	5.8	10732	21	AAI02494	Gene encoding a su
31	39.2	5.6	1635	22	AAI10594	Probe #12713 for g
32	39.2	5.6	1635	22	AAI48082	Probe #16768 used
33	39.2	5.6	1635	22	AAI08454	Probe #8445 used t
34	39.2	5.6	1973	22	AAI13556	Probe #3489 for ge
35	39.2	5.6	1973	22	AAI34918	Probe #3604 used t
36	39.2	5.6	1973	22	AAI03446	Probe #3437 used t
37	38.8	5.5	1452	21	AAI263704	Human secreted pro
38	38.6	5.5	2551	22	AAH17742	Human cDNA sequenc
39	38.4	5.5	1165	21	AAC38965	Arabidopsis thalia
40	38.4	5.5	1183	21	AAC76371	Human OREF ORF1926
41	38.4	5.5	1733	22	AA503059	Human diagnostic a
42	37.6	5.4	439	22	AAI222718	Probe #12651 for g
43	37.6	5.4	439	22	AAI48016	Probe #16702 used
44	37.6	5.4	439	22	AAI08383	Probe #8374 used t
45	37.6	5.4	1709	21	AAI15695	Human prostate can

ALIGNMENTS

RESULT	1
AAAF23528	
ID	AAF23528 standard; DNA; 700 BP.
XX	
AC	AAF23528;
XX	
DT	22-MAR-2001 (first entry)
XX	
DE	Mouse NADE DNA.
XX	
KW	Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
XX	neurogenetic disease; NF-kappaB; ds.

Sato T;

```

XX The present invention relates to a purified polypeptide capable of
CC binding neurotrophin receptor (p75-NTR). The invention is useful for
CC binding and modulating the activity of p75NTR. The peptide mediates
CC NGF-induced apoptosis, which plays an important role in neurogenetic
CC diseases. The peptide of the invention and p75NTR are useful for
CC inhibiting NF-kappaB activation in a cell or a subject, for inducing
CC caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase
CC and fragment nuclear DNA in a cell by co-expression of (1) and p75-NTR.
XX Sequence 700 BP; 177 A; 188 C; 203 G; 132 T; 0 other;

Query Match          99.1%; Score 693.6; DB 22; Length 700;
Best Local Similarity 99.4%; Pred. No. 1.2e-182;
Matches 696; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 acgagcgtctggcagcagctcgagctcctctgcgcgcgcgcgcgcgcgcgcgcgcg 60
DB 1 acgagcgtctggcagcagctcgagctcctctgcgcgcgcgcgcgcgcgcgcgcgcg 60
QY 61 aggcagcgggacagattactggaagcgcgagagtcgcagcgcgcgcgcgcgcgcg 120
DB 61 aggcagcgggacagattactggaagcgcgagagtcgcagcgcgcgcgcgcgcgcg 120
QY 121 aggactacgcccaaggatagcccgagatgacgaatgacacaggaacaaaatctcatgg 180
DB 121 aggactacgcccaaggatagcccgagatgacgaatgacacaggaacaaaatctcatgg 180
QY 181 ccaatgtccaccagaaacgaagagctgagcagcgcctctgagagtcgcagaggaacacc 240
DB 181 ccaatgtccaccagaaacgaagagctgagcagcgcctctgagagtcgcagaggaacacc 240
QY 241 gccctgtggagagagtgaggccaccagcctgctgcacaaacacacacacacacacac 300
DB 241 gccctgtggagagagtgaggccaccagcctgctgcacaaacacacacacacacacac 300
QY 301 accataaacacacacacacacacagagcgcgagcgcctgagcgcctcctaaactccgat 360
DB 301 accataaacacacacacacacacagagcgcgagcgcctgagcgcctcctaaactccgat 360
QY 361 gggccattcccaacagggcagatgaatgacgggttggtggagatggagatgatggaaa 420
DB 361 gggccattcccaacagggcagatgaatgacgggttggtggagatggagatgatggaaa 420
QY 421 ttctcatggagagatgagagatccggagaaagcttaggagatcagctgagaaatt 480
DB 421 ttctcatggagagatgagagatccggagaaagcttaggagatcagctgagaaatt 480
QY 481 gtctacgcatccttatgggggagctgtctaaacacacacacacacacacacacacac 540
DB 481 gtctacgcatccttatgggggagctgtctaaacacacacacacacacacacacacac 540
QY 541 ttatgcttgactctgggtcattccccctcgagatccatctgtgactcccgctgagccc 600
DB 541 ttatgcttgactctgggtcattccccctcgagatccatctgtgactcccgctgagccc 600
QY 601 ttctcctgcatttctcctgacatgctcttaataagaccgtttgtggtgagccctgtgtat 660
DB 601 ttctcctgcatttctcctgacatgctcttaataagaccgtttgtggtgagccctgtgtat 660
QY 661 ttccatgcatgtgccaggtggggctgtgtgtgtgcccagtgaa 700
DB 661 ttccatgcatgtgccaggtggggctgtgtgtgtgtgcccagtgaa 700

RESULT 2
ID AAF23529
XX AAF23529 standard; DNA; 891 BP.
AC AAF23529;
XX
DT 22-MAR-2001 (first entry)

```

```

XX Human NADE DNA.
DE
XX Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
KW neurogenetic disease; NF-kappaB; ds.
XX Homo sapiens.
XX WO200075278-A2.
PN 14-DEC-2000.
XX 07-JUN-2000; 2000WO-US15621.
XX 07-JUN-1999; 99US-0327750.
XX (UYCO ) UNIV COLUMBIA NEW YORK.
PA Sato T;
XX WPI; 2001-061707/07.
DR
XX New p75-neurotrophin receptor-associated cell death executor (NADE) and
PT the gene encoding NADE, useful for modulating the activity of p75NTR
PT and for detecting neurodegenerative diseases.
XX Disclosure: Fig 1; 134pp; English.
PS
XX The present invention relates to a purified polypeptide capable of
CC binding neurotrophin receptor (p75-NTR). The invention is useful for
CC binding and modulating the activity of p75NTR. The peptide mediates
CC NGF-induced apoptosis, which plays an important role in neurogenetic
CC diseases. The peptide of the invention and p75NTR are useful for
CC inhibiting NF-kappaB activation in a cell or a subject, for inducing
CC caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase
CC and fragment nuclear DNA in a cell by co-expression of (1) and p75-NTR.
XX Sequence 891 BP; 251 A; 182 C; 224 G; 234 T; 0 other;

Query Match          43.9%; Score 307.4; DB 22; Length 891;
Best Local Similarity 78.0%; Pred. NO. 1.6e-75;
Matches 428; Conservative 0; Mismatches 76; Indels 45; Gaps 3;

QY 149 aatacaccaggaacacacacacacacacacacacacacacacacacacacacacacac 208
DB 149 aatacaccaggaacacacacacacacacacacacacacacacacacacacacacacac 208
QY 209 ggagcagcctctgagagtgagacagacacacacacacacacacacacacacacacacac 268
DB 209 ggagcagcctctgagagtgagacagacacacacacacacacacacacacacacacacac 268
QY 344 ggagcagcctctgagagtgagagagacacacacacacacacacacacacacacacacac 403
DB 344 ggagcagcctctgagagtgagagagacacacacacacacacacacacacacacacacac 403
QY 269 gctgtgtcgaacacacacacacacacacacacacacacacacacacacacacacacac 328
DB 269 gctgtgtcgaacacacacacacacacacacacacacacacacacacacacacacacac 328
QY 329 ccaggctcgcgcagcctgccttaacttcgcagtgagcattcccaacacagcgagatgaatga 388
DB 329 ccaggctcgcgcagcctgccttaacttcgcagtgagcattcccaacacagcgagatgaatga 388
QY 425 acagcgtcgcgcagcctgccttaacttcgcagtgagcattcccaacacagcgagatgaatga 484
DB 425 acagcgtcgcgcagcctgccttaacttcgcagtgagcattcccaacacagcgagatgaatga 484
QY 389 cgggtgtggtgagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 448
DB 389 cgggtgtggtgagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 448
QY 485 tgggatgggtggagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 544
DB 485 tgggatgggtggagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 544
QY 449 gagaaagcttagggagctacagctagagaaattgtctacgcatccttatggggagcgtctc 508
DB 449 gagaaagcttagggagctacagctagagaaattgtctacgcatccttatggggagcgtctc 508
QY 545 aagaaagcttagggagctacagctagagaaattgtctacgcatccttatggggagcgtctc 604
DB 545 aagaaagcttagggagctacagctagagaaattgtctacgcatccttatggggagcgtctc 604
QY 509 taaccacacacacacacacacacacacacacacacacacacacacacacacacacacac 568
DB 509 taaccacacacacacacacacacacacacacacacacacacacacacacacacacacac 568
QY 605 taatcaccatgacacatcagatgagatgagatgagatgagatgagatgagatgagatgag 663
DB 605 taatcaccatgacacatcagatgagatgagatgagatgagatgagatgagatgagatgag 663
QY 569 tgagatccatactgtgactcccgctgtagccttcccttccttccttccttccttccttcct 628
DB 569 tgagatccatactgtgactcccgctgtagccttcccttccttccttccttccttccttcct 628

```


XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX OS Homo sapiens.
 XX PN EP1074617-A2.
 XX PD 07-FEB-2001.
 XX PF 28-JUL-2000; 2000EP-0116126.
 XX PR 29-JUL-1999; 99JP-0248036.
 XX PR 27-AUG-1999; 99JP-0300253.
 XX PR 11-JAN-2000; 2000JP-0118776.
 XX PR 02-MAY-2000; 2000JP-0183767.
 XX PR 09-JUN-2000; 2000JP-0241899.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX DR WPI; 2001-318749/34.
 XX PT Primer sets for synthesizing polynucleotides, particularly the 5602
 XX PT full-length cDNAs defined in the specification, and for the detection
 XX PT and/or diagnosis of the abnormality of the proteins encoded by the
 XX PT full-length cDNAs -
 XX PS Claim 1; SEQ ID 352; 2537pp + CD ROM; English.
 XX CC The present invention describes primer sets for synthesizing 5602
 XX CC full-length cDNAs defined in the specification. Where a primer set
 XX CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
 XX CC to the complementary strand of a polynucleotide which comprises one of
 XX CC the 5602 nucleotide sequences defined in the specification, where the
 XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 XX CC of an oligonucleotide comprising a sequence complementary to the
 XX CC complementary strand of a polynucleotide which comprises a 5'-end
 XX CC sequence and an oligonucleotide comprising a sequence complementary to a
 XX CC polynucleotide which comprises a 3'-end sequence, where the
 XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
 XX CC the 5'-end sequence/3'-end sequence is selected from those defined in
 XX CC the specification. The primer sets can be used in antisense therapy and
 XX CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 XX CC particularly full-length cDNAs. The primers are also useful for the
 XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
 XX CC the full-length cDNAs. The primers allow obtaining of the full-length
 XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 XX CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 XX CC represent oligonucleotides, all of which are used in the exemplification
 XX CC of the present invention.
 XX SQ Sequence 865 BP; 239 A; 177 C; 224 G; 221 T; 4 other;

Query Match 11.5%; Score 80.2; DB 22; Length 865;
 Best Local Similarity 56.2%; Pred. No. 1.5e-12;
 Matches 269; Conservative 0; Mismatches 153; Indels 57; Gaps 4;
 QY 143 gccagaaatagcaaccaggaacaaataatctcatcattgcccacccaggaacaa 202
 DB 196 gtccaaaggaggaactagcggaacaaatctcaacggggaataatgcccaacagaaacga 255
 QY 203 agagctggagcag---ccctcgagaatggacaggaacacccgcctgtggaggaggtga 259
 DB 256 aggaggggagcagcccccacgcagaaatgaagaagaatccgccttggagggggtga 315
 QY 260 gggccaccagcctgtctgaaacaaacaaacaaacaaacaaacaaacaaacaaacaa 319
 DB 316 agggcagaagcctggagggaatatca----- 341

QY 320 ccgaagagggccagctgcgcgacctgcccctaactcccgatgggcccattccccacaggca 379
 DB 342 -ggcgggggaggttagcgagctgtccctaattttcgcgatggccataacctaaggca 399
 QY 380 gatgaatgacgggtgtgggtggagatggagatgatataaggaaatttccatggagagatgag 439
 DB 400 tattgagcaca-----tgaagcgagagatgatgagaaaggtttgtaggcgagatgat 453
 QY 440 agagatcccgagagaagcttagggagctacagctgagagaattgtctacgcatccttatggg 499
 DB 454 gdaaatcaagagaagaagactagggaacagagatgagagcactatcgctctcaaa---- 509
 QY 500 ggagctgtctaaaccacacgacatcacatgatgaattctgccttacccttgacttcggctc 559
 DB 510 -----ctcctgaaacctgacaacacattatgacttttgcctcatcacttgaaatctaaa 561
 QY 560 attccccctgagatccatactgtgactcccgctgtagcccttccctcgcatcttccct 618
 DB 562 agtttctcgtaggttaattgaacactgctttacaagctgtgtattttgtgatttact 620
 RESULT 9
 AAH13750
 ID AAH13750 standard; cDNA; 1229 BP.
 XX AC AAH13750;
 XX DT 26-JUN-2001 (first entry)
 XX DE Human cDNA sequence SEQ ID NO:10656.
 XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX OS Homo sapiens.
 XX PN EP1074617-A2.
 XX PD 07-FEB-2001.
 XX PF 28-JUL-2000; 2000EP-0116126.
 XX PR 29-JUL-1999; 99JP-0248036.
 XX PR 27-AUG-1999; 99JP-0300253.
 XX PR 11-JAN-2000; 2000JP-0118776.
 XX PR 02-MAY-2000; 2000JP-0183767.
 XX PR 09-JUN-2000; 2000JP-0241899.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX DR WPI; 2001-318749/34.
 XX PT Primer sets for synthesizing polynucleotides, particularly the 5602
 XX PT full-length cDNAs defined in the specification, and for the detection
 XX PT and/or diagnosis of the abnormality of the proteins encoded by the
 XX PT full-length cDNAs -
 XX PS Claim 8; SEQ ID 10656; 2537pp + CD ROM; English.
 XX CC The present invention describes primer sets for synthesizing 5602
 XX CC full-length cDNAs defined in the specification. Where a primer set
 XX CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
 XX CC to the complementary strand of a polynucleotide which comprises one of
 XX CC the 5602 nucleotide sequences defined in the specification, where the
 XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 XX CC of an oligonucleotide comprising a sequence complementary to the
 XX CC complementary strand of a polynucleotide which comprises a 5'-end
 XX CC sequence and an oligonucleotide comprising a sequence complementary to a
 XX CC polynucleotide which comprises a 3'-end sequence, where the
 XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
 XX CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 1229 BP; 351 A; 243 C; 296 G; 339 T; 0 other;

Query Match 11.5%; Score 80.2; DB 22; Length 1229;
Best Local Similarity 56.2%; Pred. No. 1.8e-12;
Matches 269; Conservative 0; Mismatches 153; Indels 57; Gaps 4;

QY 143 gccagaatagcaaccaggaaacaaatctcatcgtgccaatgtccaccaggaaacga 202
Db 196 gtccaaagaggaactagcggaacaatctcaacggggaataatgcccaacagaaacga 255
QY 203 agagctggagcag---ccctgcagaatggacagaaacccctgtggaggagggtga 259
Db 256 agagggggcagggcccccacgcagaatgaagaataatccgccatttgggggggtga 315
QY 260 gggccaccagcctgctgcaaacacaaacaaacacacacacacacacacacacac 319
Db 316 aggcagaagcctggaggaaatatac----- 341
QY 320 ccgaagagcaggcctgcgcacccgacctccacttccgatggccattcccaacaggca 379
Db 342 --ggcggggcgagtaggcgactgtccctaatttcgatggccatacctaataaggca 399
QY 380 gatgaatcacgggttgggtggagatggagatgatgaaatgttcattggaggagatgag 439
Db 400 tattgagcaca-----tgaagcagagatgatgagaaggttgttgggcagatgat 453
QY 440 agagatccggagaagcttaggagctacagctgagaattgtctacgcattctatggg 499
Db 454 ggaatacgaagagaagactagggaacagcagatgaggcaactatagcgttccaaa---- 509
QY 500 ggagctgtctaacaccacgacatcacatgatgaattctgccttgccttgacttcggtc 559
Db 510 -----ctcctgaacctgacaacattatgactttgcttcataacctgaaactctaaa 561
QY 560 attccccctgagatcacatctgactcccgctgtagcccttccctcgcatttct 618
Db 562 agtttctgtaggttaattgaaacactgctttacaagctgtatttttggatttact 620

RESULT 10

AAC85548
ID AAC85548 standard; cDNA; 1364 BP.

XX AAC85548;

XX AAC85548;
XX 04-JUN-2001 (first entry)

DE cDNA encoding CDIFF-4, Incyte ID No. 1990956CB1.

XX Human; cell differentiation; CDIFF; agonist; antagonist; epilepsy;
KW cell proliferation; Alzheimer's disease; schizophrenia disorder;
KW arteriosclerosis; cancer; atherosclerosis; diabetes mellitus; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
XX CDS 250..612

FT /*tag= a
FT /product= "CDIFF-4"

XX WO200119860-A2.

XX 22-MAR-2001.

XX 14-SEP-2000; 2000WO-US25435.

XX 15-SEP-1999; 99US-0154140.

PR 06-DEC-1999; 99US-0169155.

XX (INCY-) INCYTE GENOMICS INC.

PI Tang YT, Hillman JL, Yue H, Reddy R, Lal P, Shah P, Azimzai Y;
PI Baughn MR, Lu DAM, Bandman O, Shih LL, Patterson C;

XX WPI; 2001-211447/21.

DR P-PSDB; AAB47126.

XX Isolated polypeptides and polynucleotides involved in cell
PT differentiation are used for treatment, prevention and diagnosis of
PT cell proliferative, developmental and neurological disorders e.g.
PT cancer and Alzheimer's disease -

XX Claim 5; Page 121; 137pp; English.

XX The sequences given in AAC85545-72 encode human polypeptides involved
CC in cell differentiation (CDIFF). CDIFF polypeptides and agonists of
CC these are used to treat a disease or condition associated with
CC decreased expression of functional CDIFF. An antagonist of CDIFF is
CC used to treat a disease or condition associated with over expression
CC of functional CDIFF. CDIFF polypeptides may be used for the treatment,
CC prevention and diagnosis of cell proliferative, developmental and
CC neurological disorders, such as Alzheimer's disease, schizophrenia
CC disorders, arteriosclerosis, cancer, atherosclerosis, diabetes mellitus
CC and epilepsy. The CDIFF-4 sequence is homologous to Mus musculus
CC REX-3. This sequence maps to chromosome 1 within the interval from
CC 152.2 to 157.4 centiMorgans, to chromosome 3 within the interval from
CC 157.4 to 158.0 centiMorgans, and to the X chromosome within the interval
CC from 104.9 to 150.3 centiMorgans.

XX Sequence 1364 BP; 411 A; 269 C; 322 G; 362 T; 0 other;

Query Match 11.5%; Score 80.2; DB 22; Length 1364;
Best Local Similarity 56.2%; Pred. No. 1.8e-12;
Matches 269; Conservative 0; Mismatches 153; Indels 57; Gaps 4;

QY 143 gccagaatagcaaccaggaaacaaatctcatcgtgccaatgtccaccaggaaacga 202

Db 255 gtccaaagaggaactagcggaacaatctcaacggggaataatgcccaacagaaacga 314

QY 203 agagctggagcag---ccctgcagaatggacagaaacccctgttggaggagggtga 259

Db 315 aggagggggcagggcccccacgcagaatgaagaataatccgccatttgggggggtga 374

QY 260 gggccaccagcctgctgcaaacacaaacacacacacacacacacacacacacacac 319

Db 375 aggcagaagcctggaggaaatatac----- 400

QY 320 ccgaagagcaggcctgcgcacccgacctccacttccgatggccattcccaacaggca 379

Db 401 --ggcggggcgagttaggcgactgtccctaatttcgatggccatacctaataaggca 458

QY 380 gatgaatcacgggttgggtggagatggagatgatgaaatgttcattggaggagatgag 439

Db 459 tattgagcaca-----tgaagcagagatgatgagaaggttgttgggcagatgat 512

QY 440 agagatccggagaagccttagggagctacagctgagaattgtctacgcattctatggg 499

Db 513 ggaatacgaagagaagactagggaacagcagatgaggcaactatgcgttccaaa---- 568

QY 500 ggagctgtctaacaccacgacatcacatgatgaattctccttatgcttgccttgccttc 559

Db 569 -----ctcctgaacctgacaacacattatgacttttgcctcattacattgaatctaaa 620

AC	AAH75810;
XX	
DT	17-OCT-2001 (first entry)
DE	Human X chromosome linked gene expression protein 14 coding sequence.
XX	
XX	Human; X chromosome linked gene expression protein 14; cancer;
KW	HIV infection; cytostatic; anti-HIV; chromosome X; ss.
XX	
OS	Homo sapiens.
XX	
PN	CN1296969-A.
XX	
PD	30-MAY-2001.
XX	
PF	23-NOV-1999; 99CN-0124078.
XX	
PR	23-NOV-1999; 99CN-0124078.
XX	(SHAN-) SHANGHAI BORONG GENE DEV CO LTD.
PA	
XX	Mao Y, Xie Y;
P1	
XX	
DR	WPI: 2001-483897/53.
DR	P-PSDB; AAG66407.
XX	
PT	Polypeptide-human X chromosome linked gene expression protein 14 and polynucleotide for coding said polypeptide -
XX	
PS	Claim 6; Page 24 (Disclosure); 31pp; Chinese.
CC	The present sequence is the coding sequence for human X chromosome linked gene expression protein 14. The protein and coding sequence are useful for treating diseases e.g. cancer and HIV infection.
XX	
SQ	Sequence 767 BP; 224 A; 143 C; 215 G; 185 T; 0 other;
Query Match 8.0%; Score 55.8; DB 22; Length 767;	
Best Local Similarity 53.7%; Pred. No. 8.4e-06;	
Matches 139; Conservative 0; Mismatches 117; Indels 3; Gaps 1	
QY	316 accaccgaagaggcaggctgcgcgacatgcccttaacttcgatggccattcccaca 375
Db	255 accgtagcggttcgcgcttaggcagccacctccgcagtatggacataatgcata 314
QY	376 ggcagatgaatgacgggttggtggagatgagatgatgaaatttcatggaggaga 435
Db	315 ggcttggagagccacaggcaaggatgagagaggagaatatggaagattggggaggag 374
QY	436 tgaagagatccgagaaagcttaggcagctcacagctgagaaaattgtctaogcatac 495
Db	375 tgaagacagctgatgaaaagctgagggaaagcagtgagtcactgcgggcagtoa 434
QY	496 tgggggagctgtctaacaccacacatccatgatgaattctgcttatgccttgacttc 555
Db	435 gcactgatccccct---caccatgacccaocatgagtttgtgcttatgcctgaatcc 491
QY	556 ggtcattccccctcctgagat 574 .
Db	492 tgatgggtttccctgaqti 510

KW	Cell cycle and proliferation protein; CCYPR; human; agonist;
KW	antagonist; gene therapy; detection; gene therapy;
KW	transgenic animal disease model; immune disorder;
KW	developmental disorder; cell signalling disorder;
KW	cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
KW	arteriosclerosis; asthma; allergy; diabetes mellitus;
KW	menstrual cycle disorder; bacterial infection; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200107471-A2.
XX	
PD	01-FEB-2001.
XX	
PF	21-JUL-2000; 2000WO-US19948.
XX	
PR	21-JUL-1999; 99US-0145075.
PR	08-SEP-1999; 99US-0153129.
PR	10-NOV-1999; 99US-0164647.
XX	
PA	(INCY-) INCYTE GENOMICS INC.
XX	
PI	Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
PI	Azimzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;
XX	
XX	WFI: 2001-112727/12.
DR	P-PSDB; AAB60474.
DR	
XX	
PT	Human cell cycle and proliferation proteins and polynucleotides are
PT	used to treat, diagnose and prevent immune, developmental and cell
PT	signaling disorders and cell proliferative disorders including cancer
XX	
XX	Claim 5; Page 181-182; 205pp; English.
XX	
CC	Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human
CC	cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.
CC	CCYPR and agonists of CCYPR are used to treat diseases or conditions
CC	associated with decreased expression of functional CCYPR, while CCYPR
CC	antagonists are used to treat diseases or conditions associated with
CC	overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
CC	to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
CC	radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
CC	compounds e.g., antibodies, oligonucleotides and proteins (receptors)
CC	that specifically bind to CCYPR, and in drug screening methods to
CC	identify compounds that modulate the activity of CCYPR. CCYPR
CC	nucleotides can be used to generate transgenic animal models of human
CC	disease, and can be used in gene therapy in target cells with genetic
CC	abnormalities with respect to the expression of CCYPR for the
CC	treatment or prevention of a disorder associated with CCYPR.
CC	Diseases which can be diagnosed, treated and prevented using CCYPR
CC	proteins, nucleic acids, agonists or antagonists include immune,
CC	developmental and cell signalling disorders, and cell proliferative
CC	disorders including cancer. Specific examples of these disorders
CC	include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
CC	diabetes mellitus, disorders of the menstrual cycle and infections
CC	caused by bacteria.
XX	
SQ	Sequence 898 BP: 250 A; 186 C; 251 G; 211 T; 0 other;

Db 528 tgaacagctgagaaagctgagggaagcaggttgagtcagtcgctgagtcac 587
||||| || | | ||||||| ||||| ||| ||||| || ||||| ||
QY 496 tggggagctgtctaacaccacagatcacccatgatgaattccttgccttgacttc 555
||||| || | ||||||| ||| || ||||| || ||||||| ||||| ||
Db 588 gcactgatcccccct---caccatgaocatcacgatgagtttgccttatgcctgaatcc 644
||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||
QY 556 ggtcattcccccttgagat 574
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 645 tgatggttccctgaagtt 663
| | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
AAC03880
ID AAC03880 standard; cDNA; 662 BP.
XX
AC AAC03880;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 3878.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
PI WPI: 2000-500381/45.
PI P-PSDB; AAC03874.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
PT
PS Claim 1; SEQ ID 3878; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dr primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 662 BP; 177 A; 144 C; 192 G; 137 T; 12 other;

Query Match 7.9%; Score 55.6; DB 21; Length 662;
Best Local Similarity 61.2%; Pred. No. 9e-06;
Matches 104; Conservative 1; Mismatches 62; Indels 3; Gaps 1;
QY 405 ggagatgatggaatgttcattgaggagatgagagatccggagaaagcttagggag 464
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 450 gaagagaatatgaaaggattggggaggagggtgagacagctgattggaagctgaggaa 509

QY 465 ctacagctgagaaaattgtctacgcacaccttatggggagctgtctaaaccacacgacac 524
||| ||||| || ||||| || | | | | | | | | | | | | | | | |
Db 510 aagcagtttgagtcatagtctcgggcagtcagcactgacccccct---caccatgacct 566
||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
QY 525 catgatgaattctgccttatgccttgacttcggtcattcccccttgagat 574
||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 567 catgatgagttttgennwatgcctgaatccctgatggtttccctaaagt 616
||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||

Search completed: March 11, 2002, 16:36:39
Job time: 3228 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2002, 16:36:39 ; Search time 144 Seconds
(without alignments)
5304.700 Million cell updates/sec

Title: us-09-327-750D-29

Perfect score: 891

Sequence: 1 accccatccccctctctat.....aataagcaatttaaaagc 891

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_ll01:*

1: /SID22/gcgdata/geneseq/geneseq/NA1980.DAT:*

2: /SID22/gcgdata/geneseq/geneseq/NA1981.DAT:*

3: /SID22/gcgdata/geneseq/geneseq/NA1982.DAT:*

4: /SID22/gcgdata/geneseq/geneseq/NA1983.DAT:*

5: /SID22/gcgdata/geneseq/geneseq/NA1984.DAT:*

6: /SID22/gcgdata/geneseq/geneseq/NA1985.DAT:*

7: /SID22/gcgdata/geneseq/geneseq/NA1986.DAT:*

8: /SID22/gcgdata/geneseq/geneseq/NA1987.DAT:*

9: /SID22/gcgdata/geneseq/geneseq/NA1988.DAT:*

10: /SID22/gcgdata/geneseq/geneseq/NA1989.DAT:*

11: /SID22/gcgdata/geneseq/geneseq/NA1990.DAT:*

12: /SID22/gcgdata/geneseq/geneseq/NA1991.DAT:*

13: /SID22/gcgdata/geneseq/geneseq/NA1992.DAT:*

14: /SID22/gcgdata/geneseq/geneseq/NA1993.DAT:*

15: /SID22/gcgdata/geneseq/geneseq/NA1994.DAT:*

16: /SID22/gcgdata/geneseq/geneseq/NA1995.DAT:*

17: /SID22/gcgdata/geneseq/geneseq/NA1996.DAT:*

18: /SID22/gcgdata/geneseq/geneseq/NA1997.DAT:*

19: /SID22/gcgdata/geneseq/geneseq/NA1998.DAT:*

20: /SID22/gcgdata/geneseq/geneseq/NA1999.DAT:*

21: /SID22/gcgdata/geneseq/geneseq/NA2000.DAT:*

22: /SID22/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	889.4	99.8	891	22	AAF23529 Human NADE DNA. H
2	564.4	63.3	917	21	AAF21748 Human breast and o
3	343	38.5	532	21	AAC01005 Human secreted pro
4	310.6	34.9	700	22	AAF23528 Mouse NADE DNA. M
5	298.8	33.5	485	21	AAC01004 Human secreted pro
6	273.8	30.7	426	22	AAH82586 Human ovarian tumo
7	160.8	18.0	169	16	AAAT21034 Human gene signatu
8	150.8	16.9	187	21	AAH87147 Rat hepatocyte car
9	141.2	15.8	865	22	AAH03517 Human cDNA clone (
10	141.2	15.8	1229	22	AAH13750 Human cDNA sequenc
11	141.2	15.8	1364	22	AAC85548 cDNA encoding CDIF

12 108.2 12.1 792 22 AAH45143 Human brain expres

13 104 11.7 898 22 AAF59611 Human cell cycle a

14 102.4 11.5 767 22 AAF75810 Human x chromosome

15 102.4 11.5 862 22 AAI58581 Human polynucleoti

16 95.4 10.7 858 22 AAI60367 Human polynucleoti

17 91.4 10.3 692 21 AAC10889 Human secreted pro

18 89.8 10.1 662 21 AAC03880 Human secreted pro

19 89.8 10.1 698 21 AAC03879 Human secreted pro

c 20 81 9.1 936 22 AAF58254 Oligonucleotide D1

c 21 81 9.1 936 22 AAF58257 Oligonucleotide D1

c 22 81 9.1 936 22 AAF58259 Oligonucleotide D2

c 23 81 9.1 936 22 AAF58262 Oligonucleotide D1

c 24 81 9.1 936 22 AAF58262 Oligonucleotide D1

c 25 74.6 8.4 1197 21 AAA72409 Human nucleic acid

26 68.2 7.7 401 21 AAC01003 Human secreted pro

27 68.2 7.7 401 21 AAC01003 Human secreted pro

28 67.4 7.6 936 22 AAF58252 Oligonucleotide D1

29 67.4 7.6 936 22 AAF58254 Oligonucleotide D1

30 67.4 7.6 936 22 AAF58257 Oligonucleotide D1

31 67.4 7.6 936 22 AAF58259 Oligonucleotide D2

32 67.4 7.6 936 22 AAF58262 Oligonucleotide D1

33 67.4 7.6 938 22 AAF58255 Oligonucleotide D1

34 66.4 7.5 128 21 AAC12100 Human secreted pro

c 35 64.8 7.3 397 22 AAI14082 Probe #4015 for ge

c 36 64.8 7.3 397 22 AAI35463 Probe #4149 used t

c 37 64.8 7.3 397 22 AAI03935 Probe #3926 used t

38 58.6 6.6 1170 22 AAH68557 Human protein HP10

39 58.6 6.6 1250 22 AAI57960 Human polynucleoti

40 58.6 6.6 1551 22 AAI59746 Human polynucleoti

41 56.4 6.3 396 21 AAH30950 Human colon cancer

42 55.4 6.2 1309 21 AAC98961 Human pancreatic c

c 43 52.8 5.9 285 22 AAI21436 Probe #11369 for g

c 44 52.8 5.9 285 22 AAI46725 Probe #15411 used

c 45 52.8 5.9 285 22 AAI07131 Probe #7122 used t

ALIGNMENTS

RESULT 1

AAF23529

ID AAF23529 standard; DNA; 891 BP.

XX AC AAF23529;

XX DT 22-MAR-2001 (first entry)

XX DE Human NADE DNA.

XX KW Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;

XX KW neurogenetic disease; NF-kappaB; ds.

XX OS Homo sapiens.

XX PN WO200075278-A2.

XX PD 14-DEC-2000.

XX PF 07-JUN-2000; 2000WO-US15621.

XX PR 07-JUN-1999; 99US-0327750.

XX PA (UYCO) UNIV COLUMBIA NEW YORK.

XX PI Sato T.

XX DR WPI; 2001-061707/07.

XX PT New p75-neurotrophin receptor-associated cell death executor (NADE) and

XX PT the gene encoding NADE, useful for modulating the activity of p75NTR

XX PS and for detecting neurodegenerative diseases -

PS Disclosure; Fig 1; 134pp; English.

XX The present invention relates to a purified polypeptide capable of
CC binding neurotrophin receptor (p75-NTR). The invention is useful for
CC binding and modulating the activity of p75NTR. The peptide mediates
CC NGF-induced apoptosis, which plays an important role in neurogenetic
CC diseases. The peptide of the invention and p75NTR are useful for
CC inhibiting NF-kappaB activation in a cell or a subject, for inducing
CC caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase
CC and fragment nuclear DNA in a cell by co-expression of (I) and p75-NTR.
XX

SQ Sequence 891 BP; 251 A; 182 C; 224 G; 234 T; 0 other:

Query Match 99.88; Score 889.4; DB 22; Length 891;
Best Local Similarity 99.9%; Pred. No. 5.9e-277;
Matches 890; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 accccatccccactctataccggtcctccatttgctgcctcaagtcctcggaag 60
DB 1 acccatccccactctataccggtcctccatttgctgcctcaagtcctcggaag 60

QY 61aatcccggaaacgaaaaatggtgggtttgggggaaaggaggttaaggggagaagctgga 120
DB 61aatcccggaaacgaaaaatggtgggtttgggggaaaggaggttaaggggagaagctgga 120

QY 121ggaggggctttaattggaggcccgtagaggacgcgcggaacttctaagtgggaaaaa 180
DB 121ggaggggctttaattggaggcccgtagaggacgcgcggaacttctaagtgggaaaaa 180

QY 181acgaaaataaaaaatccttggatcacgggctctgaaacctctgctcagagcaccgaagc 240
DB 181acgaaaataaaaaatccttggatcacgggctctgaaacctctgctcagagcaccgaagc 240

QY 241attoagtcctctccttgccctttgttgttacttggttccaagaaaaaacaccagaaaaa 300
DB 241attcagtcctctccttgccctttgttgttacttggttccaagaaaaaacaccagaaaaa 300

QY 301aaaatcctcatcgcaaatattcacaggaaaaacgaagatatgagcagcctatgcaga 360
DB 301aaaatcctcatcgcaaatattcacaggaaaaacgaagatatgagcagcctatgcaga 360

QY 361atggagaggaaacgccctttgggagagggtgaagccaccagcctcgaggaataatcgac 420
DB 361atggagaggaaacgccctttgggagagggtgaagccaccagcctcgaggaataatcgac 420

QY 421ggggacaggctcgccgacttgcctctaatttgcgatgggccaatacccaataggcagatca 480
DB 421ggggacaggctcgccgacttgcctctaatttgcgatgggccaatacccaataggcagatca 480

QY 481atgatggatgggtggagatggagatgatgaaaattattcattgagagatgagagaaa 540
DB 481atgatgggaTgggtgggagaTggagatgatgTgaaaTattcattgagagatgagagaaa 540

QY 541tcagaagaaaacttagggagctgcagttgaggaattgtctgcgtatccttatggggagc 600
DB 541tcagaagaaaacttagggagctgcagttgaggaattgtctgcgtatccttatggggagc 600

QY 601tcctcataccatgacccatcatgatgaaatttgccttatgccttgactcctgcatttcc 660
DB 601tcctcataccatgacccatcatgatgaaatttgccttatgccttgactcctgcatttcc 660

QY 661tcatgagattaactatgattcccctgtttcttttcttcttcttgcatttccctaataatgc 720
DB 661tcatgagattCaatAcgTcgattcccgctgtttcttttcttcttgcatttccctaataatgc 720

QY 721ctttactgacccgttttgcgtgaacccctaTgttatttccatgtgtcAagtgggtctctgtg 780
DB 721ctttactgacccgttttgcgtgaacccctaTgttatttccatgtgtcAagtgggtctctgtg 780

QY 781ttgcacgcttctatttgaagaattgcctttgcactcagtcgaagtttctctgcagcagtagt 840
DB 781ttgcacgcttctatttgaagaattgcctttgcactcagtcgaagtttctctgcagcagtagt 840

Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 326 ccaggaaacgaagagatggagcagcctatgcagaatggagagagagcgcctttggg 385
|||||
Db 305 ccaggaaacgaagagatggagcagcctatgcagaatggagagagcgcctttggg 364
Qy 386 aggaggtgaagccaccagcctgcaggaaatgcagcgggagcgcctgcgcacttgcgcc 445
|||||
Db 365 aggaggtgaagccaccagcctgcaggaaatgcagcgggagcgcctgcgcacttgcgcc 424
Qy 446 taatttcgatgggcccatacccaatagcagatcaatgatgggatgggtgagatggaga 505
|||||
Db 425 taatttcgatgggcccatacccaatagcagatcaatgatgggatgggtgagatggaga 484
Qy 506 tqatatgaataatcatctggaggagatgagagaaatcagaagaaacttagggagctgca 565
|||||
Db 485 tqatatgaataatcatctggaggagatgagagaaatcagaagaaacttagggagctgca 544
Qy 566 gttagggaattgtctgccttcttcttggggagcctctctaatcaccatgaccatcatga 625
|||||
Db 545 gttagggaattgtctgccttcttcttggggagcctctctaatcaccatgaccatcatga 604
Qy 626 tgaatttcgcttgccttgcactcctccatttgcattatgcagattatgcattgattccc 685
|||||
Db 605 tgaatttcgcttgccttgcactcctccatttgcattatgcagattatgcattgattccc 664
Qy 686 gctgtttctttcttctgcatttctcctaataatgcctttactgacccgttgcgtgaac 745
|||||
Db 665 gctgtttctttcttctgcatttctcctaataatgcctttactgacccgttgcgtgaac 724
Qy 746 cctatgttatttccatgtgtcaagtgggtcttctgttgcagccttctatttgaagattgc 805
|||||
Db 725 cctatgttatttccatgtgtcaagtgggtcttctgttgcagccttctatttgaagattgc 784
Qy 806 ctttgcactcagtgtaagtcttctgcagcagtagtttccaccatttgcagtgagaaattt 865
|||||
Db 785 ctttgcactcagtgtaagtcttctgcagcagtagtttccaccatttgcagtgagaaattt 844
Qy 866 aaagcctaataagcaatttaaaagc 891
|||||
Db 845 aaagcctaataagcaatttaaaagc 870

RESULT 3

AAAC01005
ID AAC01005 standard; cDNA; 532 BP.

XX
AC AAC01005;

XX
DT 06-OCT-2000 (first entry)

XX
DE Human secreted protein 5' EST, SEQ ID NO: 1003.

XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.

XX
OS Homo sapiens.

XX
PN EP1033401-A2.

XX
PD 06-SEP-2000.

XX
PF 21-FEB-2000; 2000EP-0200610.

XX
PR 26-FEB-1999; 99US-0122487.

XX
PA (GEST) GENSET.

XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX
XX WPI; 2000-500381/45.

XX
DR P-PSDB; AAG00999.

XX

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
PS Claim 1; SEQ ID 1003; 71pp + CD-ROM; English.

XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.

XX
SQ Sequence 532 BP; 151 A; 121 C; 163 G; 91 T; 6 other;

Query Match 38.5%; Score 343; DB 21; Length 532;
Best Local Similarity 98.6%; Pred. No. 6e-93;
Matches 340; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 281 agaaaaaacacagaaaaaaatctctatcgtgcaaatattccaccagaaaacgaaga 340
|||||
Db 188 agaaaaaacacagaaaaaaatctctatcgtgcaaatattccaccagaaaacgaaga 247
Qy 341 gatgagagcctatgcagaatggagagaaacaccccttggagaggtgaagggcca 400
|||||
Db 248 gatgagagcctatgcagaatggagagaaacaccccttggagaggtgaagggcca 307
Qy 401 ccagcctgcagaaatcgacggggacagctcgcgacttgccttaatttcgatgggc 460
|||||
Db 308 ccagcctgcagaaatcgacggggacagctcgcgacttgccttaatttcgatgggc 367
Qy 461 catacccaatagcagatcaatgatgggtgggtgagatgagatgataatgaaatatt 520
|||||
Db 368 catacccaatagcagatcaatgatgggtgggtgagatgagatgataatgaaatatt 427
Qy 521 catgagagatgagagaaatcagaagaaacttagggagctgcagttgagaaattgtct 580
|||||
Db 428 catggagagatgagagaaatcagaagaaacttagggagctgcagttgagaaattgtct 487
Qy 581 gcgtatccttggggagcctctctaatcaccatgaccatcatga 625
|||||
Db 488 gcgtatccttggggagcctctctaatcaccatgaccatcatga 532

RESULT 4

AAF23528
ID AAF23528 standard; DNA; 700 BP.

XX
AC AAF23528;

XX
DT 22-MAR-2001 (first entry)

XX
DE Mouse NADE DNA.

XX
KW Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
KW neurogenetic disease; NF-kappab; ds.

XX
OS Mus sp.

XX
PN WO200075278-A2.

XX
XX 14-DEC-2000.

XX
PF 07-JUN-2000; 2000WO-US15621.

XX 07-JUN-1999; 99US-0327750.
XX PA (UYCO) UNIV COLUMBIA NEW YORK.
XX Sato T;
XX WPI; 2001-061707/07.
DR New p75-neurotrophin receptor-associated cell death executor (NADE) and
PT the gene encoding NADE, useful for modulating the activity of p75NTR
PT and for detecting neurodegenerative diseases -
XX Claim 12; Fig 1; 134pp; English.
XX The present invention relates to a purified polypeptide capable of
CC binding neurotrophin receptor (p75-NTR). The invention is useful for
CC binding and modulating the activity of p75NTR. The peptide mediates
CC NGF-induced apoptosis, which plays an important role in neurogenetic
CC diseases. The peptide of the invention and p75NTR are useful for
CC inhibiting NF-kappaB activation in a cell or a subject, for inducing
CC caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase
CC and fragment nuclear DNA in a cell by co-expression of (I) and p75-NTR.
XX
SQ Sequence 700 BP; 177 A; 188 C; 203 G; 132 T; 0 other;

Query Match 34.9%; Score 310.6; DB 22; Length 700;
Best Local Similarity 78.3%; Pred. No. 3.8e-83;
Matches 430; Conservative 0; Mismatches 74; Indels 45; Gaps 3;
Qy 284 aaacaacacagaaaaaaatctcatcatgcaaaattaccacagagaaacgaagat 343
Db 149 aatagcaaccagaaaatctcatcatgcccattgtccaccaggaacgaagact 208
Qy 344 ggagcagcctatgcagatggagagagaccgcctcttgaggaggtgagggccacca 403
Db 209 ggagcagccttgagatggacagagagaccgcctcttgaggaggtgagggccacca 268
Qy 404 gcttcagagaa-----atcgacgggg 424
Db 269 gctctgcaaaacaaacaaacaaacaaacaaacaaacaaacaaacaaacaaaggg 328
Qy 425 acagctgcgcgactgcccttaatttcgatgggccatacccaatagcagatcaatga 484
Db 329 ccagctgcgcgactgcccttaatttcgatgggccatacccaatagcagatcaatga 388
Qy 485 tgggatgggtggagatggagatgatggaatattcatggagggatgagagaatcag 544
Db 389 cgggttggtggagatggagatgatggaatattcatggagggatgagagaatcag 448
Qy 545 aagaaaacttagggagctgcagtgagagaattgtcgtatccctatgggggagctctc 604
Db 449 gagaagcttagggagctacagctagagaattgtcgtatccctatgggggagctctc 508
Qy 605 taatcaccatgaccatcatgatgaatttgccttatgcttgactcgtgacttccccccc 663
Db 509 taacaccacacatcacctatgatgaatttgccttatgcttgactcgtgacttccccccc 568
Qy 664 tgagatgaatactgtgattcccgctgttttttttcttcttgcatttttcttaataatgcctt 723
Db 569 tgagatgacatactgtgactcccgctgtagccttttcttgcatttttctgacatgcctt 628
Qy 724 tactgatccgtttgctgtgaacctatgttattt-----caatgtgcaagtgggcttg 778
Db 629 taatgacccttgggtgagccttggtgtatttttccatgctatgctgaggtgggcttg 688
Qy 779 tgttgccag 787
Db 689 tgttgccag 697

AAC01004

ID AAC01004 standard; cDNA; 485 BP.

XX AAC01004;

AC AAC01004;

XX 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 1002.

DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.

OS Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GSEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX P-PSDB; AAG00998.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
CC obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
CC diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 1; SEQ ID 1002; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.

SQ Sequence 485 BP; 125 A; 113 C; 159 G; 84 T; 4 other;

Query Match 33.5%; Score 298.8; DB 21; Length 485;
Best Local Similarity 99.0%; Pred. No. 1.1e-79;
Matches 297; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 326 ccaggaacacgaagagatggagcagcctatgcagaatggaggaagacccctttggg 385

Db 186 ccaggaacacgaagagatggagcagcctatgcagaatggaggaagacccctttggg 245

Qy 386 aggggtgaaggccaccagcctcaggaatcgacgggacaggtcgcgacttgcgcc 445

Db 246 aggggtgaaggccaccagcctcaggaatcgacgggacaggtcgcgacttgcgcc 305

Qy 446 taatttcgatgggccaaccataagcagatcgaatggatgggtggagatgggaga 505

Db 306 taatttcgatgggccaaccataagcagatcgaatggatgggtggagatgggaga 365

Qy 506 tgatatggaataattcatggaggagatgagagaaatcagaagaaacttagggactgca 565

Db 366 tgatatggaataattcatggaggagatgagagaaatcagaagaaacttagggactgca 425

QY 566 gttgaggaattgtcgcgtatcttctgtggggagctctcttaataaccatgaccatcatga 625
Db :|||||
426 kttaggaattgtcgcgtatcttctgtggggakctctcttaataccatgaccatcatga 485

RESULT 6

AAH82586
ID AAH82586 standard; cDNA; 426 BP.

XX AC AAH82586;

XX DT 25-SEP-2001 (first entry)

XX DE Human ovarian tumour associated polynucleotide sequence SEQ ID NO:210.

XX KW Human; ovarian tumour; ovarian cancer; diagnosis; gene therapy;

XX KW immunogenic; vaccine; ss.

XX XX Homo sapiens.

XX OS WO200151513-A2.

XX PN 19-JUL-2001.

XX PD 16-JAN-2001; 2001WO-US01575.

XX PF 14-JAN-2000; 2000US-0176722.

XX PR (CORI-) CORIXA CORP.

XX PA Algate PA;

XX PI WPI; 2001-425866/45.

XX DR Novel ovarian tumor proteins, and nucleic acids encoding them, used to
XX PT treat and diagnose cancers, particularly ovarian cancer -

XX PS Claim 5; Page 100; 338pp; English.

XX CC AAH82377 to AAH83878 represent human ovarian tumour-associated
XX CC polynucleotide sequences which encode ovarian tumour proteins. The
XX CC ovarian tumour protein and polynucleotide sequences have cytostatic
XX CC activity, and can be used in gene therapy and vaccine production. The
XX CC ovarian tumour proteins and polynucleotides can be used to inhibit
XX CC the development of cancer, particularly ovarian cancer. They can also
XX CC be used to diagnose the onset and progression of cancer.

XX SQ Sequence 426 BP; 86 A; 106 C; 90 G; 137 T; 7 other;

Query Match 30.7%; Score 273.8; DB 22; Length 426;
Best Local Similarity 97.5%; Pred. No. 3.2e-72;
Matches 308; Conservative 0; Mismatches 5; Indels 3; Gaps 3;

QY 552 cttagggagctgacgttgagggaattgtcgtatcttctgtggggagctctctaatcac 611

Db :|||||
1 cttagggagctgcanttgagggaattgtcgtatcttctgtggggagctctctaatcac 60

QY 612 catgaccatcatgagaatttgcccttatgccttgactctccattatcatgagatta 671

Db :|||||
61 catgaccatcatgagaatttgcccttatgccttgactctccattatcatgagatta 120

QY 672 atactgtatcccgctgtttctttcttcttgcatcttccataatgcctttactgac 731

Db :|||||
121 atactgtatcccgctgtttctttcttcttgcatcttccataatgcctttactgac 179

QY 732 cgtttgctgtgaacctgtatttccatgtgtcaagtgggtcttctgtgt-tgccagctt 790

Db :|||||
180 cgtttgctgtgaacctgtatgcta-ttccatgtgtcaagtgggtcttctgtgtgtcgcagctt 238

QY 791 ctatttgaagattgcctttgcactcagtgtaagtcttctgtcagcagtagtttcaaccatt 850

Db :|||||
239 ctatttgaagattgcctttgcactcagtgtaagtcttctgtcagcagtagtttcaaccatt 298

QY 851 tgcatggaaaaatttta 866
Db :|||||
299 tgcatggaaaaatttta 314

RESULT 7

AAAT21034
ID AAT21034 standard; cDNA to mRNA; 169 BP.

XX AC AAT21034;

XX DT 03-JUL-1996 (first entry)

XX DE Human gene signature HUMGS02316.

XX KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
XX KW human; cloning; mapping; non-biased library; diagnosis; detection;
XX KW cell typing; abnormal cell function; ss.

XX OS Homo sapiens.

XX PN WO9514772-A1.

XX PD 01-JUN-1995.

XX PF 11-NOV-1994; 94WO-JP01916.

XX PR 12-NOV-1993; 93JP-0355504.

XX PA (MATS/) MATSUBARA K.

XX PA (OKUB/) OKUBO K.

XX PI Matsubara K, Okubo K;

XX DR WPI; 1995-206931/27.

XX PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
XX PT for diagnosis of abnormal cell function, by preparing cDNA that
XX PT reflects relative abundance of corresp. mRNA in specific human
XX PT tissues

XX PS Claim 1; Page 786; 2245pp; Japanese.

XX CC A single-stranded DNA (or its complementary strand or the corresp.
XX CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
XX CC given in AAT19001-T26837 and which is able to hybridise to part of
XX CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
XX CC sequences were obtained from 3'-directed cDNA libraries prepared
XX CC from various human tissues; synthesis of cDNA was initiated from the
XX CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
XX CC untranscribed sequence is unique to a particular mRNA species, almost
XX CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
XX CC is constructed so as to reflect accurately the relative abundance of
XX CC different mRNAs in the particular tissue from which it was derived.
XX CC The appearance frequency of a given GS in a cDNA library can be
XX CC determined (esp. using primers and probes derived from the GS
XX CC sequences) as a means of diagnosing abnormal cell function or for
XX CC recognising different cell types.

XX SQ Sequence 169 BP; 47 A; 30 C; 34 G; 58 T; 0 other;

Query Match 18.0%; Score 160.8; DB 16; Length 169;
Best Local Similarity 98.8%; Pred. No. 1.7e-38;
Matches 162; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 728 gatccgtttgcgtgaacctatgtttatccatgtgtcaagtgggtctgtgttccag 787

Db :|||||
1 gatccgtttgcgtgaacctatgtttatccatgtgtcaagtgggtctgtgttccag 60

QY 788 ctctcatttgaagattgcctttgcactcagtgtaagtcttctgcagcagtagtttcaccc 847
Db :|||||

Db 61 ctctattgaagattgcgtttgcactcagtgaaattctgtcagcagtagttcacc 120

Qy 848 attgcattgaaaaatttaagccaataaagcaatttaaaaaac 891

Db 121 attgcattgaaaaatttaagccaataaagcaatttaaaaaac 164

RESULT	8
AAA87147	
ID	AAA87147 standard; DNA; 187 BP.
XX	
XX	AAA87147;
XX	
XX	AC
XX	DT
XX	08-JAN-2001 (first entry)
XX	
DE	Rat hepatocyte carcinogenesis biomarker nucleic acid SEQ ID NO:71.
XX	
XX	Rat; phenobarbital; carcinogenesis marker; carcinogenesis; detection;
KW	identification; carcinogenic; probe; primer; ds.
XX	
XX	Rattus norvegicus.
OS	
XX	WO200044902-A2.
PN	
XX	03-AUG-2000.
PD	
XX	
XX	28-JAN-2000; 2000WO-US00503.
PF	
XX	
XX	29-JAN-1999; 99US-0118078.
PR	
XX	
XX	(SEAR) SEARLE & CO G D.
PA	
XX	
PI	Bunch RT, Curtis SW, Rodi CP, Morris DL;
XX	
XX	WPI; 2000-505977/45.
XX	
XX	New nucleic acid encoding a carcinogenic biomarker, induced by
PT	phenobarbital treatment of rat hepatocytes, useful for identifying
PT	carcinogenic compounds -
XX	
XX	Claim 1; Page 73; 240pp; English.
PS	
XX	
XX	AAA87080 to AAA87656 represent nucleic acid sequences (N1) encoding a
CC	carcinogenesis biomarkers. The carcinogenesis biomarkers are induced
CC	treating rat hepatocytes with phenobarbital. The nucleic acids are
CC	useful for identifying carcinogenic compounds. The nucleic acid molec
CC	can be used to derive probes and/or primers for detecting or inducing
CC	carcinogenesis, respectively.
XX	
XX	Sequence 187 BP; 39 A; 48 C; 40 G; 60 T; 0 other:

RESULT 9
 AAH03517
 ID AAH03517 standard; cDNA: 865 BP.
 XX
 XX AC
 XX AAH03517;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 XX
 DE Human cDNA clone (5'-primer) SEQ ID NO:352.
 XX
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 KW
 XX Homo sapiens.
 OS
 XX
 XX EPI074617-A2.
 PN
 XX
 XX 07-FEB-2001.
 PD
 XX
 XX 28-JUL-2000; 2000EP-0116126.
 PF
 XX
 XX 29-JUL-1999; 99JP-0248036.
 PR
 XX 27-AUG-1999; 99JP-0300253.
 PR
 XX 11-JAN-2000; 2000JP-0118776.
 PR
 XX 02-MAY-2000; 2000JP-0183767.
 PR
 XX 09-JUN-2000; 2000JP-0241899.
 XX
 XX (HELI-) HELIX RES INST.
 PA
 XX
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 XX WPI: 2001-318749/34.
 XX
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 XX Claim 1: SEQ ID 352: 2537pp + CD ROM; English.
 PS

Db 423 ggattgggagggggtgagacagctgatgaaaaagctgagggaaagcagttgagtcata 482
 QY 577 gtcgcgtatccttatggggagctctctaatcaccatgaccatcatgatgaatttggc 636
 Db 483 gtcgaggcagtcagcactgacccccc---tcaccatgaccatcatgatgatttggc 539
 QY 637 ttatgctgactcctgcccattatcatgagatttaatactggtatcccgctgtttctt 696
 Db 540 ttatgcccgaactcgtggtttccctaaagtattac--ggaaacagaccctgcttt 597
 QY 697 ttctctgcatcttctaataatgcttttactgatccggtttgtgtgaaacctatattt 756
 Db 598 cgaattacatgttcatgatgagccctgtgtgaaacctttacctgtcactgttt---- 653
 QY 757 tccatgtgcaagtggtctgtgtgcccagcttctatttgagattgcttt 809
 Db 654 -----acgtgggtcctctattaccagcttctaattgaattgtgttt 696

RESULT 13
 AAF59611
 ID AAF59611 standard; cDNA; 898 BP.
 XX
 AC AAF59611;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Human cell cycle and proliferation protein CCYPR-22 cDNA, SEQ ID NO:76.
 XX
 KW Cell cycle and proliferation protein; CCYPR; human; agonist;
 KW antagonist; gene therapy; detection; gene therapy;
 KW transgenic animal disease model; immune disorder;
 KW developmental disorder; cell signalling disorder;
 KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
 KW arteriosclerosis; asthma; allergy; diabetes mellitus;
 KW menstrual cycle disorder; bacterial infection; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200107471-A2.
 XX
 PD 01-FEB-2001.
 XX
 PF 21-JUL-2000; 2000WO-US19948.
 XX
 PR 21-JUL-1999; 99US-0145075.
 PR 08-SEP-1999; 99US-0153129.
 PR 10-NOV-1999; 99US-0164647.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
 PI Azimzal Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P;
 XX
 DR WPI; 2001-1127/27/12.
 DR P-PSDB; AAB60474.
 XX
 PT Human cell cycle and proliferation proteins and polynucleotides are
 PT used to treat, diagnose and prevent immune, developmental and cell
 PT signalling disorders and cell proliferative disorders including cancer -
 XX
 PS Claim 5; Page 181-182; 205pp; English.
 XX
 CC Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human
 CC cell cycle and proliferation proteins (CCYPR), AAB60453-AAB60506.
 CC CCYPR and agonists of CCYPR are used to treat diseases or conditions
 CC associated with decreased expression of functional CCYPR, while CCYPR
 CC antagonists are used to treat diseases or conditions associated with
 CC overexpression of functional CCYPR. Monoclonal or polyclonal antibodies
 CC to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or
 CC radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect
 CC compounds e.g., antibodies, oligonucleotides and proteins (receptors)
 CC that specifically bind to CCYPR, and in drug screening methods to

CC identify compounds that modulate the activity of CCYPR. CCYPR
 CC nucleotides can be used to generate transgenic animal cells of human
 CC disease, and can be used in gene therapy in target cells with genetic
 CC abnormalities with respect to the expression of CCYPR for the
 CC treatment or prevention of a disorder associated with CCYPR.
 CC diseases which can be diagnosed, treated and prevented using CCYPR
 CC proteins, nucleic acids, agonists or antagonists include immune,
 CC developmental and cell signalling disorders, and cell proliferative
 CC disorders including cancer. Specific examples of these disorders
 CC include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies,
 CC diabetes mellitus, disorders of the menstrual cycle and infections
 CC caused by bacteria.
 XX
 SQ Sequence 898 BP; 250 A; 186 C; 251 G; 211 T; 0 other;
 Query Match 11.7%; Score 104; DB 22; Length 898;
 Best Local Similarity 55.4%; Pred. No. 4.9e-21;
 Matches 346; Conservative 0; Mismatches 240; Indels 38; Gaps 6;

QY 283 aaaaacacacagaaaaaaatctcatctgcaaaatattcaccagagaaacgaagaga 342
 Db 264 aagaggaaacgagctttaacaatctcatctcgtggaataatgcaacaggaaaatgatgaa 323
 QY 343 tggagcagcctatgcagaatggagaggaaga-----ccgccctttgggaggag 390
 Db 324 aagatgaaagagcaagttgctaataaaggaggagcccttggccctacatttgaaattga 383
 QY 391 gtgaaggccacacgactgcaggaaa---tcgacggggacagagctgcgacttgcacctc 447
 Db 384 gtgaatactgtgctcctagagaaacgtaggcggttcgcgcttaggcagccatctgc 443
 QY 448 atttcgattggccatacccaatagggacagatcaatgatggatgggtgagatgagatg 507
 Db 444 agtatagatggacataatgcataggcttggagagccacaggaaggtgagagaggaga 503
 QY 508 atatgaaatattcatggaggagatgagagaaatcagaagaaacttagggagctgcagt 567
 Db 504 atatgaaagattggggaggaggtgagacagctgatgaaaagctgagggaaaagcagt 563
 QY 568 tgaggaattgctgcgtatcccttatggggagctctctataatcaccatgaccatcatg 627
 Db 564 tgagtcatagtctgcggcagtcagcactgatccccct---caccatgaccatcagatg 620
 QY 628 aattttgcttgccttgactcctgccttcttatcatgagattatcatgattactgtatccgc 687
 Db 621 agttttccttgccttgccttgccttgccttgccttgccttgccttgccttgccttgc 680
 QY 688 tgttttcttcttgccttgccttgccttgccttgccttgccttgccttgccttgccttgc 747
 Db 681 ttcctaaacttaca-----catttgggtgacctttgtcgttaaacg 722
 QY 748 tatgtatttccatgtgtcaagtggtcttgccttgccttgccttgccttgccttgccttgc 807
 Db 723 ttttgatgttacctatttcttgccttgccttgccttgccttgccttgccttgccttgc 782
 QY 808 ttg-cactcagtgtaagtttctgcagcagtag-tttcaccatttgcattggaataattt 865
 Db 783 ttgaccaggtttgaagtttctgcagcagtagttttacatttgcattggaataattt 842
 QY 866 aaagcccaataaagcaatttaaaaa 889
 Db 843 tcattatattgtgaagttaata 866

RESULT 14
 AAF75810
 ID AAF75810 standard; cDNA; 767 BP.
 XX
 AC AAF75810;
 XX
 DT 17-OCT-2001 (first entry)
 XX

THIS PAGE BLANK (USPTO)

```
110 CAACCAAGAAAAATCAAGAAAAAGGAGCAAGTTGCTAATAATAAAGGGAGCCCT 159
34 euAlaLeuProLeuAspAlaGlyClnTyrCysValProArgGlyAsnArg 50
|||||
160 TGGCCCTCCCTTTGGATGCTGGTGAATACATGTCCTAGAGGAATCGT 209
51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67
210 AGGCGGTTCGCGCTAGGACGCCCATCTGCAGTATAGATGGGATATGAT 259
67 tHisArgLeuGlyClnProGlnAlaArgMetArgGluGluAsnMetGluA 84
260 GCATAGGCTTTGGAGAACCAAGCAAGGATGAGAGAGCAAGATATGGAAA 309
84 rgileGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
310 GGATTGGGAGGAGGTGAGACACTGATGGAAGAGCTGAGGAAAGACAG 359
101 LeuSerHisSerLeuArgAlaValSerThrAspProProHisHisAspHi 117
360 TTGAGTCATAGTCTGCGGCGAGTCAGCACTGACCCCTCACCATGACCA 409
117 sHisAspGluPheCysLeuMetPro 125
410 TCATGATGAGTTTGGCTTATGCCCC 434
```

```
seq_name: gb_est1:AI929703
seq_documentation_block: 576 bp mRNA EST 23-AUG-1999
LOCUS AI929703
DEFINITION au63f04.y1 Schneider fetal brain 00004: Homo sapiens cDNA clone
IMAGE: 2519455 5' similar to TR:088858 O88858 REX-3.; mRNA
sequence.
ACCESSION AI929703
VERSION AI929703.1 GI:5665667
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 576)
Hillier, L., Allen, M., Bowles, L., Duboue, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Lennon, G., Marra, M., Martin
J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reversed clone: similarity on wrong strand
Seq primer: -40RP from Gibco
High quality sequence stop: 444.
Location/Qualifiers
1..576
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2519455"
/clone.lib="Schneider fetal brain 00004"
/sex="male"
/tissue_type="frontal lobe"
/lab_host="DH10B"
/dev_stage="5 months post-conception"
/notes="Organ: brain; Vector: pBluescript SK (Stratagene);
Site_1: SstI; Site_2: XhoI; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence:
```

```
FEATURES
source
```

```
5'-GAGAGAGAGAGAGCTCAAGGATCCTTAATTAATTAATCCCCCCCCCC-3'
and 3' adaptor sequence:
5'-GAGAGAGAGAGCTCGAGCTTTT-3'. The library was
size-selected for >0.5 kb inserts and has an average
insert size estimated at 1.2 kb. This library was
constructed using the CAP-trapper method for full-length
enrichment and has not undergone amplification. Library
was constructed by Dr. Claudio Schneider (LNCIB-Area
Science Park, Trieste, Italy)."
BASE COUNT 154 a 141 c 176 g 104 t 1 others
ORIGIN
alignment_scores:
Quality: 646.00 Length: 125
Ratio: 5.252 Gaps: 0
Percent Similarity: 98.400 Percent Identity: 97.600
alignment_block:
US-09-327-750D-32 x AI929703
Align seg 1/1 to: AI929703 from: 1 to: 576
1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
202 ATGGAGTCCAAAGAGAAACGAGCAGTAAACAGTCTCAGCATGGAATGC 251
17 aAsnGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProL 34
|||||
252 CAACCAAGAAAAATCAAGAAAAAGGAGCAAGTTGCTAATAAAGGGAGCCCT 301
34 euAlaLeuProLeuAspAlaGlyClnTyrCysValProArgGlyAsnArg 50
|||||
302 TGGCCCTCCCTTTGGATGCTGGTGAATACATGTCCTAGAGGAATCGT 351
51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67
|||||
352 AGGCGGTTCGCGCTAGGACGCCCATCTGCAGTATAGATGGGATATGAT 401
67 tHisArgLeuGlyClnProGlnAlaArgMetArgGluGluAsnMetGluA 84
|||||
402 GCATAGGCTTTGGAGAACCAAGCAAGGATGAGAGAGCAAGATATGGAAA 451
84 rgileGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
|||||
452 GGATTGGGAGGAGGTGAGACACTGATGGAAGAGCTGAGGAAAGACAG 501
101 LeuSerHisSerLeuArgAlaValSerThrAspProProHisHisAspHi 117
|||||
502 TTGAGTCATAGTCTGCGGCGAGTCAGCACTGACCCCTCACCATGACCA 551
117 sHisAspGluPheCysLeuMetPro 125
552 TCATGATGAGTTTGGCTTATGCCCC 576
seq_name: gb_est2:BG715659
seq_documentation_block: 734 bp mRNA EST 08-MAY-2001
LOCUS BG715659
DEFINITION 602676924F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4799476 5',
mRNA sequence.
ACCESSION BG715659
VERSION BG715659.1 GI:13994846
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 734)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
```

520 TCATGATGAGTTTTCCTTATGCC 544

seq_name: gb_est1:AL526247

seq_documentation_block: 794 bp mRNA EST 13-FEB-2001
 LOCUS AL526247 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC016Y001 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL526247
 VERSION AL526247.1 GI:12789740
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 794)
 AUTHORS Li.W.B., Gruber.C., Jesse.J. and Polayes.D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1..794
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DC016Y001"
 /clone_lib="LTI_NFL003_NBC3"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed
 by Life Technologies. Contact : Feng Liang Life
 Technologies, a division of Invitrogen 9800 Medical Center
 Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
 8371 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 236 a 163 c 213 g 180 t 2 others

ORIGIN

alignment_scores:
 Quality: 653.00 Length: 125
 Ratio: 5.266 Gaps: 0
 Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:

US-09-327-750D-32 x AL526247

Align seg 1/1 to: AL526247 from: 1 to: 794

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetClnuAsnAl 17
 148 ATGGAGTCCAAAGAGAAACAGCAGTAAACAGTCTCAGCATGGAAATGC 197
 17 aAsnGlnGluAsnGluGluLysGlnValAlaAsnLysGlyGluProL 34
 198 CAACCAAGAAATGAAGAAAGAGGAGCAAGTGTCTATAAAGGGGAGCCCT 247
 34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
 248 TGGCCCTCCCTTTGGATGCTGGTGAATACTGTGTGCTAGAGGAATCGT 297
 51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67
 298 AGCGGTTCCGGTTCAGGAGCCCATCCCTGCAGATAGATAGATATGAT 347
 67 tHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84

348 GCATAGGCTTGGAGAACCCACAGGATGAGAGAAGACAATATGGAA 397

84 rgIleGlyGluGluValArgGlnLeuMetClnuLysLeuArgGluLysGln 100

398 GGATTGGGGAGGAGGTGAGACAGCTGATGGAAAGCTGAGGGAAGCAG 447

101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 117

448 TTGAGTCATAGTCTGGGGCAGTCAGACATGACCCCTCACCATGACCA 497

117 sHisAspGluPheCysLeuMetPro 125

498 TCATGATGAGTTTTCCTTATGCC 522

seq_name: gb_est1:AV702643

seq_documentation_block: 683 bp mRNA EST 08-OCT-2000
 LOCUS AV702643 ADB Homo sapiens cDNA clone ADBDBE11 5', mRNA sequence.
 DEFINITION AV702643
 ACCESSION AV702643
 VERSION AV702643.1 GI:10718973
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao
 H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,
 Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu
 G., Hu, R., Chen, J., Chen, Z. and Han, Z.

TITLE

Homo sapiens cDNA ADB clones

JOURNAL

Unpublished (2000)

COMMENT

Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919(ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.

FEATURES

source

1..683
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="ADBDBE11"
 /clone_lib="ADB"
 /tissue_type="Adrenal gland"
 /dev_stage="Adult"
 /lab_host="SOLR"
 /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
 XhoI"

BASE COUNT 208 a 128 c 184 g 161 t 2 others

ORIGIN

alignment_scores:
 Quality: 650.00 Length: 125
 Ratio: 5.242 Gaps: 0
 Percent Similarity: 99.200 Percent Identity: 97.600

alignment_block:

US-09-327-750D-32 x AV702643

Align seg 1/1 to: AV702643 from: 1 to: 683

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetClnuAsnAl 17
 60 ATGGAGTCCAAAGAGAAACAGCAGTAAACAGTCTCAGCATGGAAATGC 109
 17 aAsnGlnGluAsnGluGluLysGlnValAlaAsnLysGlyGluProL 34


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4933012"
/clone_lib="NCI_CGAP_Brn67"
/tissue_type="anaplastic oligodendroglioma with lp/19q loss"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: brain; Vector: pCMV-SPOrt6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      198 a   138 c   191 g   172 t
ORIGIN
alignment_scores:
  Quality: 653.00      Length: 125
  Ratio: 5.266        Gaps: 0
  Percent Similarity: 99.200      Percent Identity: 98.400
alignment_block:
US-09-327-750D-32 x BG820179
Align seg 1/1 to: BG820179 from: 1 to: 699
1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
94 ATGGAGTCCAAAGAGAAACGACGAGTAACAGTCTCAGCATGGAAATGC 143
17 aAsnGlnGluAsnGluGluGluGlnValAlaAsnLysGlyGluProL 34
144 CAACCAAGAAATGAAGAAAGGAGCAAGTTGCTAATAAAGGGAGCCCT 193
34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
194 TGGCCCTCCCTTTGGATGCTGGTGAATACTGTGCTAGAGGAAATCGT 243
51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTyrAspIleMe 67
244 AGCGGTTCCGGCTTAGGAGCCCATCTGCAGTATAGATGGATATGAT 293
67 thisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
294 GCATAGGCTTGGAGAACCAACGACGCAAGGATGAGAGAGAGAAATATG 343
84 rGileGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
344 GGATGGGAGGAGGTGAGACAGCTGATGGAAGCTGAGGAAAGAGCAG 393
101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 117
394 TTGAGTCATAGTCTGGGGCAGTCAGCACTGACCCCTCCACCATGACCA 443
117 shiAspGluPheCysLeuMetPro 125
444 TCATGATGAGTTTTCCTTATGCC 468
seq_name: gb_est1:BE790774

seq_documentation_block:
LOCUS      BE790774      783 bp      mRNA
DEFINITION 601581956F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3936613 5', mRNA sequence.
ACCESSION  BE790774
VERSION    BE790774.1 GI:10211972
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 783)
AUTHORS  NIH-MGC http://mgc.nci.nih.gov/.
TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)

```

JOURNAL COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCFD/DPF
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LICM780 row: O column: 14
 High quality sequence start: 3
 High quality sequence stop: 780.

FEATURES

source

```

1..783
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3936613"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: POTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGCAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 214 a 169 c 219 g 181 t
 ORIGIN

alignment_scores:

Quality: 653.00 Length: 125
 Ratio: 5.266 Gaps: 0
 Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:

US-09-327-750D-32 x BE790774

Align seg 1/1 to: BE790774 from: 1 to: 783

```

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
170 ATGGAGTCCAAAGAGAAACGACGAGTAACAGTCTCAGCATGGAAATGC 219
17 aAsnGlnGluAsnGluGluGlnValAlaAsnLysGlyGluProL 34
220 CAACCAAGAAATGAAGAAAGGAGCAAGTTGCTAATAAAGGGAGCCCT 269
34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
270 TGGCCCTCCCTTTGGATGCTGGTGAATACTGTGCTAGAGGAAATCGT 319
51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTyrAspIleMe 67
320 AGCGGTTCCGGCTTAGGAGCCCATCTGCAGTATAGATGGATATGAT 369
67 thisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
370 GCATAGGCTTGGAGAACCAACGACGCAAGGATGAGAGAGAGAAATATG 419
84 rGileGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
420 GGATGGGAGGAGGTGAGACAGCTGATGGAAGCTGAGGAAAGAGCAG 469
101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 117
470 TTGAGTCATAGTCTGGGGCAGTCAGCACTGACCCCTCCACCATGACCA 519
117 shiAspGluPheCysLeuMetPro 125

```

60 ATGGAGTCCAAAGAGAAACGACGAGTAAACAGTCTCAGCATGGAATAATGC 109
 17 asnGlnGluAsnGluGluLysGluGlnValAlaAsnLysGluProL 34
 |||||
 110 CAACCAAGAAATGAAGAAAGGACGAGTTCCTAATAAAGGGAGCCCT 159
 |||||
 34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
 |||||
 160 TGGCCCTCCCTTTGGATGCTGCTGAATACTGTGTGCTAGAGAAATCGT 209
 |||||
 51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67
 |||||
 210 AGCGCGTTCCCGCTTAGGCGAGCCATCTCTGCAGTATAGATGATGAT 259
 |||||
 67 thisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGlu 84
 |||||
 260 GCATAGGCTTGGAGAACACGACGAGGATGAGAGAAGATATGGAAA 309
 |||||
 84 rgileGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
 |||||
 310 GCATTGGGAGGAGGTGACACAGCTGTGGAAGAGCTGAGGGAAGACAG 359
 |||||
 101 LeuSerHisSerLeuArgAlaValSerThrAspProProHisHisAspHi 117
 |||||
 360 TTGAGTCTATAGTCTCGGGCAGCTCAGCACTGACCCCTCACCATGACCA 409
 |||||
 117 shiAspGluPheCysLeuMetPro 125
 |||||
 410 TCATGATGAGTTTTCCTTATGCC 434

seq_name: gb_est1:AL526209

seq_documentation_block: 695 bp mRNA EST 13-FEB-2001
 LOCUS AL526209 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC016Y001 3
 DEFINITION AL526209 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC016Y001 3

prime, mRNA sequence.

ACCESSION AL526209

VERSION AL526209.1 GI:12789702

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization.

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. .695

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DC016Y001"

/clone_lib="LTI_NFL003_NBC3"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed

by Life Technologies. Contact : Feng Liang Life

Technologies, a division of Invitrogen 9800 Medical Center

Drive Rockville, Maryland 20850, USA Fax : (1) 301 610

8371 Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

158 a 196 c 154 g 183 t 2 others

BASE COUNT
 ORIGIN

alignment_scores:

Quality: 653.00 Length: 125

Ratio: 5.266 Gaps: 0

Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:

US-09-327-750D-32 x AL526209/rev ..

Align seg 1/1 to reverse of: AL526209 from: 1 to: 695

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17

|||||

546 ATGGAGTCCAAAGAGAAACGACGAGTAAACAGTCTCAGCATGGAATAATGC 497

|||||

17 asnGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProL 34

|||||

496 CAACCAAGAAATGAAGAAAGGACGAGTTCCTAATAAAGGGAGCCCT 447

|||||

34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50

|||||

446 TGGCCCTCCCTTTGGATGCTGCTGAATACTGTGCTAGAGAAATCGT 397

|||||

51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67

|||||

396 AGCGCGTTCCCGCTTAGGCGAGCCATCTCTGCAGTATAGATGATGAT 347

|||||

67 thisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGlu 84

|||||

346 GCATAGGCTTGGAGAACACGACGAGGATGAGAGAAGATATGGAAA 297

|||||

84 rgileGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100

|||||

296 GCATTGGGAGGAGGTGACACAGCTGTGGAAGAGCTGAGGGAAGACAG 247

|||||

101 LeuSerHisSerLeuArgAlaValSerThrAspProProHisHisAspHi 117

|||||

246 TTGAGTCTATAGTCTCGGGCAGCTCAGCACTGACCCCTCACCATGACCA 197

|||||

117 shiAspGluPheCysLeuMetPro 125

|||||

196 TCATGATGAGTTTTCCTTATGCC 172

|||||

seq_name: gb_est2:BG820179

seq_documentation_block: 699 bp mRNA EST 22-MAY-2001

LOCUS BG820179 699 bp mRNA EST 22-MAY-2001

DEFINITION 602782222F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4933012

5', mRNA sequence.

ACCESSION BG820179

VERSION BG820179.1 GI:14167766

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 699)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: David N. Louis, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAMI0859 row: 1 column: 05

High quality sequence stop: 695.

FEATURES
 source
 1. .699

```
101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 117
473 TTGAGTCATAGTCTGGCGGACGTCAGCACTGACCCCTCACCATGACCA 522
117 sHisAspGluPheCysLeuMetPro 125          EST      08-OCT-2000
523 TCATGATGAGTTTGGCTTATGCCC 547

seq_name: gb_est1:AV702285
seq_documentation_block:
LOCUS   AV702285      690 bp      mRNA          EST          08-OCT-2000
DEFINITION   AV702285 ADB Homo sapiens cDNA clone ADBBXE07 5', mRNA sequence.
ACCESSION   AV702285
VERSION     AV702285.1 GI:10718615
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 690)
AUTHORS     Peng, Y., Song, H., Huang, Q., Gu, Y., Yang, Y., Gao, G., Xiao
            , H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,
            Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu
            , G., Hu, R., Chen, J., Chen, Z. and Han, Z.
TITLE       Homo sapiens cDNA ADB clones
JOURNAL     Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES    Location/Qualifiers
             source
               1..690
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="ADBBXE07"
               /clone_lib="ADB"
               /tissue_type="Adrenal gland"
               /dev_stage="Adult"
               /lab_host="SOLR"
               /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
               XhoI"
BASE COUNT  210 a 132 c 180 g 166 t 2 others
ORIGIN

alignment_scores:
  Quality: 653.00      Length: 125
  Ratio: 5.266        Gaps: 0
  Percent Similarity: 99.200      Percent Identity: 98.400

alignment_block:
US-09-327-750D-32 x AV702285
..
Align seg 1/1 to: AV702285 from: 1 to: 690

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
60 ATGGAGTCCAAAGAGAAACGACGAGTAAACAGTCTCAGCATGAAATGC 109
17 aAsnGlnGluAsnGluLysGluGlnValAlaAsnLysGlyGluProL 34
110 CAACCAAGAAATGAAGAAAGGAGGACAGTTGCTAATAAAGGGAGCCCT 159
34 euAlaLeuProLeuAspAlaGlyGlyTyrCysValProArgGlyAsnArg 50
160 TGCCCTCCCTTTGGATGCTGGTGACTACTGTGCTGCTAGAGGAATCGT 209
51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67
```

```
117 sHisAspGluPheCysLeuMetPro 125
410 TCATGATGAGTTTGGCTTATGCCC 434

seq_name: gb_est1:AV705808
seq_documentation_block:
LOCUS   AV705808      691 bp      mRNA          EST          09-OCT-2000
DEFINITION   AV705808 ADB Homo sapiens cDNA clone ADBCRH03 5', mRNA sequence.
ACCESSION   AV705808
VERSION     AV705808.1 GI:10723103
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 691)
AUTHORS     Peng, Y., Song, H., Huang, Q., Gu, Y., Yang, Y., Gao, G., Xiao
            , H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,
            Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu
            , G., Hu, R., Chen, J., Chen, Z. and Han, Z.
TITLE       Homo sapiens cDNA ADB clones
JOURNAL     Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES    Location/Qualifiers
             source
               1..691
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="ADBCRH03"
               /clone_lib="ADB"
               /tissue_type="Adrenal gland"
               /dev_stage="Adult"
               /lab_host="SOLR"
               /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
               XhoI"
BASE COUNT  211 a 133 c 181 g 162 t 4 others
ORIGIN

alignment_scores:
  Quality: 653.00      Length: 125
  Ratio: 5.266        Gaps: 0
  Percent Similarity: 99.200      Percent Identity: 98.400

alignment_block:
US-09-327-750D-32 x AV705808
..
Align seg 1/1 to: AV705808 from: 1 to: 691

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
```

http://image.llnl.gov
 Plate: LIAM10003 row: d column: 14
 High quality sequence stop: 658.
 Location/Qualifiers
 source
 1..658

FEATURES

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4361053"
 /clone_lib="NIH_MGC_84"
 /tissue_type="adrenal cortex carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site_1:
 NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT
 primed. Average insert size 1.229 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 205 a 124 c 164 g 165 t
 ORIGIN

alignment_scores:
 Quality: 653.00 Length: 125
 Ratio: 5.266 Gaps: 0
 Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:

US-09-327-750D-32 x BF971303 ..

Align seg 1/1 to: BF971303 from: 1 to: 658

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
 |||||

24 ATGGAGTCCAAAGAGAAACGACGAGTAACAGCTCTCAGCATGGAAATGC 73

17 aasnGlnGluAsnGluGluGluGluGluGluGluGluGluGluGluGlu 34

74 CAACCAAGAAATGAAGAAAGGACGACAGTTGCTAATAAAGGGAGCCCT 123

34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50

124 TGGCCCTCCCTTTGGATGCTGGTGAATCTGTGCTAGAGGAATCGT 173

51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67

174 AGCGGTTCGCGGTTAGGCAGCCCATCTGCAGTATAGATGGATATGAT 223

67 thisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84

224 GCATAGGCTTGGAGAACCCACAGGCAAGGATGAGAGAAGATATGGAAA 273

84 rgileGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100

274 GGATGGGAGAGGTGAGACAGCTGTGGAAGAGCTGAGGGAAGAGCAG 323

101 LeuSerHisSerLeuArgAlaValSerThrAspProProHisHisAspHi 117

324 TTGAGTCATAGTCTCGGGCAGTCAGCACTGACCCCTCACCATGACCA 373

117 shiAspGluPheCysLeuMetPro 125

374 TCATGATGAGTTTGCCTTATGCC 398

seq_name: gb_est2:BG714823

seq_documentation_block:

LOCUS BG714823 674 bp mRNA 08-MAY-2001
 DEFINITION 602677146F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4799747 5',

mRNA sequence.
 ACCESSION BG714823

VERSION BG714823.1 GI:13993754

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 674)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LIAM10689 row: k column: 12
 High quality sequence stop: 672.

FEATURES

source

Location/Qualifiers

1..674

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4799747"

/clone_lib="NIH_MGC_96"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/note="Organ: brain; Vector: pBluescriptR (modified

pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag

); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',

size-selected for average insert size 2.3 kb and

normalized to R0T 5. This is a primary library enriched

for full-length clones and constructed using the

Cap-trapper method (Carninci, in preparation). Library

constructed by M. Brownstein (NIH/NHGRI, National

Institutes of Health). Note: this is a NIH_MGC Library."

175 a 156 c 201 g 142 t

BASE COUNT

ORIGIN

alignment_scores:

Quality: 653.00 Length: 125
 Ratio: 5.266 Gaps: 0
 Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:

US-09-327-750D-32 x BG714823 ..

Align seg 1/1 to: BG714823 from: 1 to: 674

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17

173 ATGGAGTCCAAAGAGAAACGACGAGTAACAGCTCTCAGCATGGAAATGC 222

17 aasnGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProL 34

223 CAACCAAGAAATGAAGAAAGGACGACAGTTGCTAATAAAGGGAGCCCT 272

34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50

273 TGGCCCTCCCTTTGGATGCTGGTGAATCTGTGCTAGAGGAATCGT 322

51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67

323 AGCGGTTCGCGGTTAGGCAGCCCATCTGCAGTATAGATGGATATGAT 372

67 thisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84

373 GCATAGGCTTGGAGAACCCACAGGCAAGGATGAGAGAAGATATGGAAA 422

84 rgileGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100

423 GGATGGGAGAGGTGAGACAGCTGATGGAAAGCTGAGGAAAGAGCAG 472

```

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
143 ATGGAGTCCAAAGAGAAACGAGCAGTAAACAGTCTCAGCATGGAATGC 192
|||||
17 aasnGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProL 34
|||||
193 CAACCAAGAAATGAAGAAAGGAGCAAGTTGCTAATAAAGGGAGCCCT 242
|||||
34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
|||||
243 TGGCCCTCCCTTTGGATGCTGGTGAATCTGTGCTAGAGGAAATCGT 292
|||||
51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67
|||||
293 AGCGCGTTCCGCTTAGGCAGCCCATCTGCAGTATAGATGGATATGAT 342
|||||
67 tHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
|||||
343 GCATAGGCTTGGAGAACACAGCAGCAAGGATGAGAGAGAGAAATATGAAA 392
|||||
84 rgIleGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
|||||
393 GGATGGGAGGAGGTGAGACAGCTGATGAAAGCTGAGGGAAGCAG 442
|||||
101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 117
|||||
443 TTCAGTCATAGTCTGGGCGAGTCAGACACTGACCCCTCACCATGACCA 492
|||||
117 sHisAspGluPheCysLeuMetPro 125
|||||
493 TCATGATGAGTTTGCCTTATGCCC 517
|||||

```

seq_name: gb_est2:BG472167

```

seq_documentation_block: 634 bp mRNA EST 21-MAR-2001
LOCUS BG472167
DEFINITION 602513894F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4645775 5',
mRNA sequence.
ACCESSION BG472167
VERSION BG472167.1 GI:13404541
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LICM1419 row: k column: 24
High quality sequence stop: 634.
Location/Qualifiers
1. 634
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4645775"
/clone_lib="NIH_MGC_16"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pOTB7; Site:1; XhoI: Site:2;
EcoRI: cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGCAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit

```

FEATURES
source

(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library.

```

BASE COUNT 187 a 119 c 166 g 162 t
ORIGIN

alignment_scores:
Quality: 653.00 Length: 125
Ratio: 5.266 Gaps: 0
Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:
US-09-327-750D-32 x BG472167
Align seg 1/1 to: BG472167 from: 1 to: 634

```

```

1 MetGluSerLysGluLysArgAlaValAsnSerLeuSerMetGluAsnAl 17
|||||
30 ATGGAGTCCAAAGAGAAACGAGCAGTAAACAGTCTCAGCATGGAATGC 79
|||||
17 aasnGlnGluAsnGluGluLysGluGlnValAlaAsnLysGlyGluProL 34
|||||
80 CAACCAAGAAATGAAGAAAGGAGCAAGTTGCTAATAAAGGGAGCCCT 129
|||||
34 euAlaLeuProLeuAspAlaGlyGluTyrCysValProArgGlyAsnArg 50
|||||
130 TGGCCCTCCCTTTGGATGCTGGTGAATCTGTGCTAGAGGAAATCGT 179
|||||
51 ArgArgPheProValArgGlnProIleLeuGlnTyrArgTrpAspIleMe 67
|||||
180 AGCGCGTTCCGCTTAGGCAGCCCATCTGCAGTATAGATGGATATGAT 229
|||||
67 tHisArgLeuGlyGluProGlnAlaArgMetArgGluGluAsnMetGluA 84
|||||
230 GCATAGGCTTGGAGAACACAGCAGCAAGGATGAGAGAGAGAAATATGAAA 279
|||||
84 rgIleGlyGluGluValArgGlnLeuMetGluLysLeuArgGluLysGln 100
|||||
280 GGATGGGAGGAGGTGAGACAGCTGATGAAAGCTGAGGGAAGCAG 329
|||||
101 LeuSerHisSerLeuArgAlaValSerThrAspProHisHisAspHi 117
|||||
330 TTCAGTCATAGTCTGGGCGAGTCAGACACTGACCCCTCACCATGACCA 379
|||||
117 sHisAspGluPheCysLeuMetPro 125
|||||
380 TCATGATGAGTTTGCCTTATGCCC 404
|||||

```

seq_name: gb_est2:BF971303

```

seq_documentation_block: 658 bp mRNA EST 22-JAN-2001
LOCUS BF971303
DEFINITION 602273150F1 NIH_MGC_84 Homo sapiens cDNA clone IMAGE:4361053 5',
mRNA sequence.
ACCESSION BF971303
VERSION BF971303.1 GI:12338518
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 658)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

OM of: US-09-327-750D-32 to: EST:* out_format : pfs

Date: Mar 11, 2002 2:16 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet.p2n.model -DEV=x1h
-O=/cgn2.1/USPTO.spool/US09327750/runat_11032002_101153_20308/app_query.fasta_1.1472
-DB=EST -OFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DLOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500
-MINLEN=0 -MAXLEN=2000000000 -USER=US09327750_ECGNL_5054
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-32

Query length: 125

Database: EST:*

Database sequences: 11351937

Database length: 1077921985

Search time (sec): 4085.940000

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	Documentation
gb_est2:BG655001	+	653.00	1221.08	7.5e-59	547	! BG655001 lb44f06.y1 HR85 islet
gb_est2:BG472167	+	653.00	1219.74	8.9e-59	634	! BG472167 602513894F1 NIH_MGC_96
gb_est2:BF971303	+	653.00	1219.40	9.3e-59	658	! BF971303 602273150F1 NIH_MGC_84
gb_est2:BG714823	+	653.00	1219.18	9.6e-59	674	! BG714823 6026771146F1 NIH_MGC_96
gb_est1:AV702285	+	653.00	1218.97	9.8e-59	690	! AV702285 AV702285 ADB Homo sapi
gb_est1:AV705808	+	653.00	1218.96	9.9e-59	691	! AV705808 AV705808 ADB Homo sapi
gb_est1:AL526209	-	653.00	1218.91	9.9e-59	695	! AL526209 AL526209 LTI_NFL003_NH
gb_est2:BG820179	+	653.00	1218.82	1.0e-58	699	! BG820179 602782222F1 NCI_CGAP_9
gb_est2:BE790774	+	653.00	1217.82	1.1e-58	783	! BE790774 601581956F1 NIH_MGC_7
gb_est1:AL526247	+	653.00	1217.70	1.2e-58	794	! AL526247 AL526247 LTI_NFL003_NH
gb_est1:AV702643	+	650.00	1213.43	2.0e-58	683	! AV702643 AV702643 ADB Homo sapi
gb_est1:AI929703	+	646.00	1207.46	4.3e-58	576	! AI929703 au63f04.y1 Schneider f
gb_est2:BG715659	+	646.00	1205.26	5.7e-58	734	! BG715659 602676924F1 NIH_MGC_96
gb_est2:BG709325	+	637.00	1187.25	5.8e-57	828	! BG709325 602673345F1 NIH_MGC_96
gb_est2:AW160832	+	631.00	1180.01	1.5e-56	531	! AW160832 au76b12.y1 Schneider f
gb_est2:BF346478	+	630.00	1176.22	2.4e-56	656	! BF346478 602020342F1 NCI_CGAP_9
gb_est1:AV727777	+	627.00	1171.87	4.1e-56	569	! AV727777 AV727777 HTC Homo sapi
gb_est1:AW953922	+	627.00	1171.28	4.5e-56	607	! AW953922 EST365887 MAGE resequ
gb_est2:BG701861	+	626.00	1168.97	6.0e-56	783	! BG701861 602683214F1 NIH_MGC_96
gb_est1:AV726527	+	626.00	1164.93	1.0e-55	993	! AV726527 AV726527 HTC Homo sapi
gb_est2:BF966650	-	625.00	1165.09	9.9e-56	794	! BF966650 602286982T1 NIH_MGC_96
gb_est2:BG714687	+	607.00	1132.41	6.5e-54	700	! BG714687 602677046F1 NIH_MGC_96
gb_est1:BE314909	+	602.50	1125.45	1.6e-53	594	! BE314909 601140074F1 NIH_MGC_9
gb_est1:AL520931	-	602.50	1123.74	2.0e-53	717	! AL520931 AL520931 LTI_NFL004_NH
gb_est2:BE889313	+	602.50	1123.07	2.2e-53	772	! BE889313 601513201F1 NIH_MGC_71
gb_est2:BG709503	+	602.50	1122.63	2.3e-53	810	! BG709503 602674781F1 NIH_MGC_96
gb_est2:BG707398	+	602.50	1122.35	2.4e-53	835	! BG707398 602672810F1 NIH_MGC_96
gb_est2:BF038787	+	602.50	1120.34	3.1e-53	1042	! BF038787 601462119F1 NIH_MGC_6
gb_est2:BG285825	+	600.50	1117.46	4.4e-53	946	! BG285825 602380917F1 NIH_MGC_93
gb_est2:BG036775	+	600.00	1118.04	4.1e-53	801	! BG036775 602286982F1 NIH_MGC_93
gb_est2:CI5948	+	598.00	1118.78	3.7e-53	488	! CI5948 CI5948 Clontech human ad
gb_est2:BF348593	+	598.00	1113.14	7.7e-53	908	! BF348593 602019889F1 NCI_CGAP_9
gb_est1:AW024587	-	597.00	1115.16	6.0e-53	591	! AW024587 wu79g08.x1 NCI_CGAP_K1
gb_est2:BG714974	+	595.50	1110.47	1.1e-52	726	! BG714974 602677263F1 NIH_MGC_96
gb_est2:BF967675	-	595.50	1109.67	1.2e-52	793	! BF967675 602287383T1 NIH_MGC_96
gb_est2:BG707734	+	594.50	1109.40	1.2e-52	817	! BG707734 602671255F1 NIH_MGC_96
gb_est1:AA160657	+	594.50	1109.81	1.3e-52	639	! AA160657 zq49f12.rl Stratagene
gb_est2:W48778	+	592.00	1107.91	1.5e-52	467	! W48778 zc44c03.rl Soares_senes
gb_est2:BG705843	+	591.50	1101.93	3.3e-52	813	! BG705843 602669339F1 NIH_MGC_96
gb_est1:AI291270	+	589.50	1102.39	3.1e-52	511	! AI291270 qm16s06.x1 NCI_CGAP_LU

gb_est1:AI291126 + 587.50 1096.19 6.8e-52 669 ! AI291126 qm15f02.x1 NCI_CGAP
gb_est2:BF237433 + 587.50 1095.63 7.3e-52 711 ! BF237433 601842108F1 NIH_MGC
gb_est1:BE266012 + 586.50 1095.14 7.8e-52 610 ! BE266012 601194731F1 NIH_MGC
gb_est1:BE733217 + 584.50 1087.88 2.0e-51 897 ! BE733217 601568452F1 NIH_MGC
gb_est2:BF967457 + 581.50 1083.65 3.4e-51 769 ! BF967457 602287383F1 NIH_MGC

seq_name: gb_est2:BG655001

seq_documentation_block:
LOCUS BG655001 547 bp mRNA EST 05-JUL-2001
DEFINITION lb44f06.y1 HR85 islet Homo sapiens cDNA 5' similar to TR:Q9WT28
Q9WT28 BRAIN EXPRESSED X-LINKED PROTEIN 2. ; mRNA sequence.

ACCESSION BG655001

VERSION BG655001.1 GI:13792410

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 547)

REFERENCE

AUTHORS

Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wyllie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R., Williams,T.,
Jackson,Y. and Bowers,Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Other_ESTs: lb44f06.x1

JOURNAL

COMMENT

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For Information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@wustl.edu)

Seq primer: -40RP from Gibco

High quality sequence stop: 447.

FEATURES

Source

1..547
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"

/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; CDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

BASE COUNT 152 a 120 c 174 g 101 t

ORIGIN

alignment_scores:

Quality: 653.00 Length: 125
Ratio: 5.266 Gaps: 0
Percent Similarity: 99.200 Percent Identity: 98.400

alignment_block:

US-09-327-750D-32 x BG655001 ..

Align seg 1/1 to: BG655001 from: 1 to: 547